A functional missense variant in ITIH3 affects protein expression and neurodevelopment and confers schizophrenia risk in Han Chinese population

Kaiqin Li, Yifan Li, Junyang Wang, Yongxia Huo, Di Huang, Shiwu Li, Jiewei Liu, Xiaoyan Li, Rong Liu, Xiaogang Chen, Yong-Gang Yao, Ceshi Chen, Xiao Xiao, Ming Li, Xiong-Jian Luo

PII: S1673-8527(20)30082-5
DOI: https://doi.org/10.1016/j.jgg.2020.04.001
Reference: JGG 770

To appear in: Journal of Genetics and Genomics

Received Date: 23 February 2020
Revised Date: 30 March 2020
Accepted Date: 2 April 2020


This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Copyright © 2020, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and Genetics Society of China. Published by Elsevier Limited and Science Press. All rights reserved.
XJL conceived and designed the study. KQL performed most of the experiments, including genotyping, vectors construction, western blotting, isolation and culture of mouse NSCs, proliferation, migration and differentiation of NSCs, and qPCR. YFL, JYW and DH performed DNA extraction and genotyping assays. SWL participated in isolation and characterization of mouse NSCs. YXH, JWL and XYL conducted the genetic association, meta-analysis, RNA-seq, GO and KEGG analyses. RL and CSC carried out the protein degradation assays. XGC, YGY, XX, ML and XJL contributed to this work in sample collection, DNA extraction, genotyping assays, results interpretation and manuscript writing. XJL oversaw the project. KQL and XJL drafted the first version of the manuscript. All authors revised the manuscript critically and approved the final version.
A functional missense variant in ITIH3 affects protein expression and neurodevelopment and confers schizophrenia risk in Han Chinese population

Kaiqin Li\textsuperscript{a,b}, Yifan Li\textsuperscript{a,b}, Junyang Wang\textsuperscript{a,b}, Yongxia Huo\textsuperscript{a}, Di Huang\textsuperscript{a}, Shiwu Li\textsuperscript{a,b}, Jiewei Liu\textsuperscript{a}, Xiaoyan Li\textsuperscript{a,b}, Rong Liu\textsuperscript{a,c}, Xiaogang Chen\textsuperscript{d}, Yong-Gang Yao\textsuperscript{a,b,c,f}, Ceshi Chen\textsuperscript{a}, Xiao Xiao\textsuperscript{a}, Ming Li\textsuperscript{a,b,c,f,*} & Xiong-Jian Luo\textsuperscript{a,b,c,e,*}

\textsuperscript{a}Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

\textsuperscript{b}Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650204, China

\textsuperscript{c}Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China

\textsuperscript{d}Institute of Mental Health, National Clinical Research Center for Mental Health Disorders and National Technology Institute of Psychiatry, The Second Xiangya Hospital, Central South University, Changsha 410011, China

\textsuperscript{e}KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

\textsuperscript{f}CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, 200031, China
*To whom correspondence should be addressed; Key Laboratory of Animal Models and Human Disease Mechanisms, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China; Tel: +86-871-68125413, Fax: +86-871-68125413, E-mail: luoxiongjian@mail.kiz.ac.cn (XJL), or wudaliming@163.com (ML)
The psychiatric genomics consortium (PGC) has recently identified 10 potential functional coding variants for schizophrenia. However, how these coding variants confer schizophrenia risk remains largely unknown. Here we investigate the associations between 8 potential functional coding variants identified by PGC and schizophrenia in a large Han Chinese sample (N=4,022 cases and 9,270 controls). Among the 8 tested SNPs, rs3617 (a missense variant, p.K315Q in the ITIH3 gene) showed genome-wide significant association with schizophrenia in Han Chinese population ($P = 8.36 \times 10^{-16}$), with the same risk allele as in PGC. Interestingly, rs3617 locates in a genomic region that is highly evolutionarily-conserved, and its schizophrenia risk allele (C allele) was associated with lower ITIH3 mRNA and protein expression. Intriguingly, mouse neural stem cells stably overexpressing ITIH3 with different alleles of rs3617 exhibited significant differences in proliferation, migration and differentiation, suggesting the impact of rs3617 on neurodevelopment. Subsequent transcriptome analysis found that the differentially expressed genes in neural stem cells stably overexpressing different alleles of rs3617 were significantly enriched in schizophrenia-related pathways, including cell adhesion, synapse assembly, MAPK and PI3K-AKT pathways. Our study provides convergent lines of evidence suggesting that rs3617 in ITIH3 likely affects protein function and neurodevelopment and thereby confers risk of schizophrenia.

Keywords: Schizophrenia, ITIH3, missense variant, association, neurodevelopment, neural stem cells
1. **Introduction**

Schizophrenia (SZ) is a severe, complex, and chronic psychiatric disorder that affects ~0.5-1% of the world’s population (Saha et al., 2005). The lifetime prevalence of schizophrenia has remained relatively high, and it was estimated that about 23.6 million people were affected by schizophrenia in 2013 (Global Burden of Disease Study 2013 Collaborators, 2015). The clinical manifestations of schizophrenia include positive symptoms (such as delusions and hallucinations), negative symptoms (e.g., impaired motivation and social withdraw) and cognitive impairments (such as poorer cognitive performance) (Owen et al., 2016). The symptoms of schizophrenia usually begin in young adulthood (between age 16 and 30), and can last for many years. Schizophrenia also leads to substantial mortality linked with suicide (Palmer et al., 2005; Saha et al., 2007; Oakley et al., 2018), and patients suffer from greater comorbidity rates of other mental health problems (Buckley et al., 2009) and cardiovascular disease (Goff et al., 2005). Therefore, schizophrenia imposes enormous burden on public well-being.

Though schizophrenia has become a major threat to global health, currently there are few effective approaches to successfully prevent or cure this disease. A major reason for the therapeutic quandary is that we know little about the causes of schizophrenia. To date, the etiology and pathophysiology of schizophrenia remain largely unclear. Nevertheless, accumulating evidence suggest that both genetic and environmental factors are involved in its pathogenesis (Walker et al., 2004). The heritability of schizophrenia was estimated to be around 0.8 (Sullivan et al., 2003), suggesting a key role of the inherited variants in the disease. To uncover the genetic basis of schizophrenia, numerous genetic linkage and association studies have been performed and multiple candidate variants and genes were identified (Lewis et al., 2003; Allen et al., 2008; Ng et
The advent of genome-wide association studies (GWASs) provides an unprecedented opportunity to dissect the genetic architecture of schizophrenia. In 2008, O'Donovan et al. reported the first genome-wide significant schizophrenia risk gene ZNF804A (O'Donovan et al., 2008). Since then, many GWASs have been conducted in different continental populations and multiple risk variants (or loci) have been identified (Shi et al., 2011; Yue et al., 2011; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Li et al., 2017; Yu et al., 2017; Pardinas et al., 2018). In 2014, PGC identified 108 genome-wide significant schizophrenia risk loci through analyzing 36,989 cases and 113,075 controls (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). For each of the reported risk loci, PGC identified potential causal SNPs through defining a credible causal set of SNPs (the possibility that these credible causal set of SNPs contain the true causal variants is 99%) (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). While most of the potential causal SNPs predicted by PGC were located in non-coding regions, known missense variants within 10 loci were considered to be potentially causal (i.e., the reported association signal from each of these 10 loci was credibly attributable to a known non-synonymous SNP).

Clearly, these non-synonymous SNPs can provide insights into the genetic mechanisms and pathogenesis of schizophrenia, yet their underlying functions and mechanisms remain undefined.

To examine the roles of these 10 potential non-synonymous polymorphisms in schizophrenia, we performed genetic association studies and functional characterization experiments in this study. We firstly tested the associations between 8 potential non-synonymous SNPs identified by PGC (rs2955365 was failed to genotype, and rs13107325 is monomorphic in Chinese population) and schizophrenia in a large Han Chinese sample (n = 4022 cases and 9270 controls), and...
demonstrated that rs3617 (p.K315Q) in the ninth exon of \textit{ITIH3} was significantly associated with schizophrenia at the genome-wide level in Han Chinese population. We then investigated the impact of rs3617 on \textit{ITIH3} mRNA and protein levels, as well as proliferation, migration and differentiation of mouse neural stem cells (mNSC) overexpressing ITIH3 mutants with different amino acids at rs3617 site. We further carried out transcriptome analysis to identify the genes differentially expressed in mNSC stably transfected with different alleles of rs3617. Our study provides convergent lines of evidence that support rs3617 as a potential causal variant for schizophrenia.

2. Results

2.1. rs3617 shows genome-wide significant association with schizophrenia in Han Chinese population

For each of the risk loci identified by PGC2 (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), the authors defined a credible causal set of SNPs. Most of the potential causal variants were in non-coding regions, except that 10 non-synonymous SNPs were considered responsible for the association signals of these loci. We hypothesized that these non-synonymous SNPs might also be associated with schizophrenia in independent populations if they were causal, and therefore investigated the associations between these 10 non-synonymous SNPs and schizophrenia in a large Han Chinese sample (4022 cases and 9270 controls). Notably, rs13107325 is monomorphic in Han Chinese population according to 1000 Genomes dataset (Abecasis et al., 2010), and rs2955365 failed to be genotyped. We thus successfully genotyped 8 non-synonymous SNPs (Table 1). The overall genotyping call rate was 99.7%. We performed
Hardy-Weinberg equilibrium tests and found that none of the genotyped SNPs deviated from Hardy-Weinberg equilibrium in controls (Table S1). Single-site association analysis showed that the coding variant rs3617 in *ITIH3* was significantly associated with schizophrenia in our sample (*P*=8.36×10^{-16}), even after Bonferroni correction for multiple testing (corrected *P*=6.69×10^{-15}). Also, the schizophrenia risk allele at rs3617 is consistently across Chinese and European populations (Pardinas et al., 2018) (Table S2), suggesting that it is likely a true risk SNP for schizophrenia.

To further validate the association between rs3617 and schizophrenia, we conducted a meta-analysis through combining samples from a recent study (40,675 cases and 64,643 controls) (Pardinas et al., 2018) and those of the current study. The meta-analysis (a total of 44,697 cases and 73,913 controls) showed that rs3617 was strongly associated with schizophrenia (*P*=1.35×10^{-19}) (Table S3). Given that rs3617 is a missense variant and a potential causal variant predicted by PGC2, our validation in an independent population (i.e., Chinese sample) further supports that rs3617 is a potentially causal risk variant for schizophrenia.

### 2.2. rs3617 is located in an evolutionary highly conserved region

Rs3617 locates in the ninth exon (Fig. 1A) of *ITIH3*. The ITIH proteins have two conserved domains, the VIT (vault protein inter-alpha-trypsin) domain and vWFA (von Willebrand factor type A) domain (Himmelfarb et al., 2004). We performed domain prediction using Prosite (https://prosite.expasy.org/) and found that rs3617 was located in the vWFA domain (Fig. 1B). Through interacting with integrins, collagens and extracellular matrix proteins, the vWFA domain plays important roles in cell adhesion (Colombatti and Bonaldo, 1991; Whittaker and Hynes,
Alignment of protein sequence from multiple vertebrate species suggests that rs3617 is in a highly conserved region (Fig. 1C), indicating vital biological roles of the proteins encoded by the sequence containing rs3617. Intriguingly, while the corresponding amino acid of the codon with rs3617 (the 315th amino acid) is lysine in most species (Fig. 1C), a different amino acid, glutamine, has emerged in humans due to the derived allele of rs3617. We further explored whether rs3617 has an effect on 3D structure of ITIH3 by using SWISS-MODEL (https://swissmodel.expasy.org/). The 3D structure prediction suggested that rs3617 did not exert obvious effect on ITIH3 structure (Fig. 1D and E).

2.3. LD patterns of the 8 studied exonic SNPs in CEU and CHB populations

We compared the overall LD structure across the genomic regions containing the 8 tested non-synonymous SNPs in the main PGC population (Europeans, CEU) and Chinese population (CHB). LD comparisons mainly based on pairwise LD values between the studied exonic SNP and its surrounding variants and the LD blocks (i.e., if the exonic SNP is located in a specific LD block or in recombination hotspot in European and Chinese populations) showed that 3 SNPs (including rs3617, rs20551, rs3176443) had different LD patterns in CEU and CHB populations (Figs. S1–S3). However, the other 5 SNPs (rs4584886, rs133335, rs950169, rs2288920 and rs3774729) showed similar LD patterns in CEU and CHB populations (Figs. S4–S8).

2.4. The risk allele (C allele) of rs3617 is associated with lower ITIH3 protein and mRNA levels

Considering that rs3617 is significantly associated with schizophrenia and locates in an evolutionary conserved region, it is possible that rs3617 affects schizophrenia risk by modulating
the pivotal biological function of ITIH3. We therefore investigated the impact of rs3617 on ITIH3 protein abundance. Expression vectors containing different alleles of rs3617 were constructed and GFP was fused to the C terminus of ITIH3. These expression vectors were transiently transfected into HEK293 cells and ITIH3 protein was quantified through western blotting after 24, 48 and 72 h from the time of transfection. We found that the schizophrenia risk C allele consistently led to a significantly lower ITIH3 protein level at different time points after transfection compared with the A allele ($P < 0.001$, Fig. 2A–D). To further validate these results, we constructed ITIH3 expression vector through tagging Flag to the N terminus of ITIH3. Again, we found that cells carrying the C allele had significantly lower protein levels of ITIH3 compared with those carrying the A allele ($P<0.05$, Fig. 2E and F). We also explored the impact of different rs3617 alleles on ITIH3 protein levels in mNSCs stably overexpressing the gene, and consistently, the C allele of rs3617 was associated with lower ITIH3 protein expression in mNSCs (Fig. 2G and H).

Theoretically, abundance of ITIH3 protein is regulated by the rates of its synthesis and degradation, and we therefore explored whether p.K315Q affects ITIH3 stability. By performing CHX chase assays (Buchanan et al., 2016), we found that the half-life of different ITIH3 isoforms (with the 315th amino acid to be either lysine or glutamine) were similar (Fig. S9), indicating that p.K315Q did not affect ITIH3 protein stability. Therefore, different alleles of rs3617 likely affect the synthesis of ITIH3 protein.

We also explored the effect of rs3617 on ITIH3 mRNA expression using qPCR and the results showed that rs3617 significantly affected mRNA levels of ITIH3 (Fig. S10) in both HEK293 cells (Fig. S10A) and mNSCs (Fig. S10B). Taken together, rs3617 affects ITIH3 expression at both protein and mRNA levels.
2.5. Association between rs3617 and gene expression in human brains

We explored the association between rs3617 and gene expression in human brain using GTEx and LIBD datasets. Rs3617 was associated with the expression of GLYCTK, GNL3 and PPM1M in GTEx (Table S4). In the LIBD dataset, rs3617 was associated with the expression of GNL3 ($P=7.89\times10^{-11}$) and NEK4 ($P=1.89\times10^{-6}$) (Table S5). No significant association between rs3617 and ITIH3 expression was observed in either dataset. We noticed that mRNA levels of ITIH3 were relatively low in human brain tissues (Fig. S11), which may explain why no significant association between rs3617 and ITIH3 expression was observed in brain eQTL datasets. In addition, considering the high LD between rs3617 and other genetic variants (Fig. S12), it is also possible that the associations between rs3617 and other genes were resulted from the signals of its linkage disequilibrium (LD) SNPs.

2.6. C allele of rs3617 promotes the proliferation of mNSCs compared with A allele

The neurodevelopmental hypothesis of schizophrenia posits that schizophrenia originates from abnormal brain development (Owen et al., 2011), and has gained support from accumulating evidence. For example, Wash et al. found that gene-disrupting rare structural variants identified in schizophrenia cases were enriched in neurodevelopmental pathways (Walsh et al., 2008). Functional studies of schizophrenia risk genes (such as DISC1 and RELN) also further confirmed the involvement of abnormal neurodevelopment in the disease (Mao et al., 2009; Ishizuka et al., 2011; Senturk et al., 2011). We showed that the risk allele of rs3617 was associated with lower ITIH3 expression. To further explore the potential functional consequence of rs3617, we
investigated the effect of rs3617 on neurodevelopment. We isolated the mNSCs from embryonic brains (E13.5) as previously described (Yang et al., 2018) and confirmed that mNSCs were successfully isolated and purified by immunofluorescence co-labeling of the well-characterized NSC markers (including SOX2, NESTIN and PAX6) (Fig. 3A–E). We transfected the mNSCs with ITIH3 expression vectors (carrying different alleles at rs3617) to generate mNSCs stably overexpressing ITIH3. Proliferation of these cells was then evaluated through the BrdU incorporation experiments. Intriguingly, mNSCs carrying the schizophrenia risk C allele at rs3617 exhibited significantly higher proliferation rates compared with those carrying the non-risk A allele ($P < 0.001$, Fig. 3F and G). Further cell proliferation experiment using CCK-8 also showed consistent results ($P < 0.01$, Fig. 3H), which determines cell viability through reduction reaction (WST-8 is reduced by dehydrogenases and the amount of the reduction products (formazan) is directly proportional to the number of living cells). We have also transiently transfected ITIH3 expression vectors into HEK293 cells, and tested the effects of different alleles at rs3617 on cell proliferation after 48 h. Again, both BrdU and CCK-8 assays suggested that transfection of vectors carrying C allele at rs3617 resulted in significantly increase of cell proliferation rate (Fig. 3I–K). Taken together, rs3617 likely affects neurodevelopment through modulating NSC proliferation.

2.7. C allele of rs3617 inhibits the migration of mNSCs compared with A allele

We further investigated the effects of C and A alleles of rs3617 on migration of mNSCs. The cells were isolated and stably transfected with vectors carrying either C or A allele at rs3617, and then cultured in proliferation medium to generate neurospheres. Neurospheres with similar diameters were seeded into laminin pre-coated plates and cultured in differentiation medium for 24 h. The
migration distances of the seeded neuropheres were then measured. We found that the migration
distances of mNSCs stably overexpressing schizophrenia risk C allele at rs3617 were significantly
shorter than those stably overexpressing A allele ($P < 0.05$) (Fig. 4). Therefore, rs3617 might also
affect neurodevelopment by modulating NSC migration.

2.8. rs3617 affects the differentiation of mNSCs

Neural differentiation, an essential process in neurodevelopment, has also been found to be
affected by schizophrenia risk genes (Yang et al., 2018; Chen et al., 2019). We therefore also
explored whether different alleles of rs3617 exerted different impacts on neural differentiation.

Briefly, mNSCs stably transfected with either C or A allele of rs3617 were differentiated into
neurons and glial cells and stained for the neuronal markers (TUJ1 and MAP2) and the glial
marker (GFAP). The ratio of GFAP (a marker for glial cells), MAP2 (a marker for mature neurons)
and TUJ1 (a marker for newly generated immature post-mitotic neurons) positive cells were
calculated and compared to reveal the differentiation ability of the mNSCs carrying different
alleles at rs3617. Indeed, significantly different differentiation capabilities were seen in cells with
different genotypes at rs3617 (Fig. 5). Compared with mNSCs stably transfected with A allele of
rs3716, the differentiation of mNSCs with the C allele into astrocytes (marked with GFAP, a
marker for astrocytes) was significantly impaired ($P<0.001$) (Fig. 5A and B). By contrast,
significant increases in the ratios of MAP2 (Fig. 5C and D) and TUJ1 (Fig. 5E and F) positive
cells were observed in mNSCs carrying the C allele compared with those stably transfected with
the A allele. These results confirmed the allelic effects of rs3617 on neural differentiation (i.e., the
ratios of glial and neuronal cells).
2.9. rs3617 does not affect ITIH3 protein localization

We then examined if different amino acids resulted from rs3617 affected ITIH3 protein localization. However, immunofluorescence showed that rs3617 did not affect the subcellular localization of ITIH3 in either HEK293 cells (Fig. S13) or mNSCs (Fig. S14). Therefore, rs3617 likely affected ITIH3 through mechanisms other than altering protein localization. In addition, to exclude the potential influences of endogenous ITIH3 on our results, we performed RT-qPCR and confirmed that no obvious endogenous ITIH3 expression was detected in HEK293 cells (Fig. S15A). In mNSCs, Itih3 expression was also very low as evaluated by both RT-qPCR (Fig. S15B) and RNA-seq. Specifically, the average expression value (counts from RNA-Seq data) of the well-characterized NSC marker Sox2 was 43,283 in mNSCs, while that of Itih3 was only 95. Taken together, endogenous ITIH3 should not affect our results significantly.

2.10. Dosage-dependent effect of ITIH3 on proliferation

Although the effect of rs3617 on cell proliferation was convincing, it was not clear whether this effect was attributed to the altered expression of the ITIH3 protein expression level, as the constructs of different rs3617 alleles might differ in their activities and thereby confound the results. We therefore conducted a dosage dependent assay, and confirmed that the ITIH3 expression level was proportional to the amount of vectors transfected C allele of rs3617 (Fig. S16A). Interestingly, we found that the proliferation rate of HEK293 cells was decreased as the transfected amount of ITIH3 expression vector increased (Fig. S16B). This data is consistent with the results presented in Fig. 3, suggesting that the observed proliferation difference was likely due
to the altered ITIH3 protein levels. Nevertheless, whether rs3617 affects the activity of ITIH3 remains to be defined.

2.11. rs3617 regulates schizophrenia-related biological pathways

Our above data suggests that rs3617 confers schizophrenia risk through affecting ITIH3 function and neurodevelopment. To further investigate the potential mechanisms underlying these functional impact of rs3617, RNA-seq analysis and subsequent qPCR validation were carried out. We identified 669 genes that were differentially expressed between mNSCs stably transfected with C allele of rs3617 and those carrying the A allele (Fig. 6A). The top 30 most significantly altered genes are showed in Fig. 6B. Five genes (including Peg10, Sgce, Dbx2, Rasgrp3 and Syt1) were selected from these top 30 genes for qPCR validation (Fig. 6B), as their pivotal roles in the brain have been consistently reported before. For example, Dbx2 is involved in neurogenesis, and Syt1 has a role in maintaining the function of synapses (Ullrich et al., 1994; Ma et al., 2011). SGCE mutations were associated with psychiatric disorders (Peall et al., 2015), and Rasgrp3 is involved in cell migration (Randhawa et al., 2011). Consistent with the RNA-seq results, qPCR showed that Peg10, Sgce and Dbx2 were significantly down-regulated in mNSCs stably transfected with the C allele of rs3617 compared with those transfected with the A allele (Fig. 6C–E); while Rasgrp3 and Syt1 were significantly up-regulated in mNSCs carrying the C allele at rs3617 (Fig. 6F–G). We then performed GO analysis and found that the differentially expressed genes were significantly enriched in cell adhesion, synapse organization, positive regulation of cell migration, gliogenesis and synapse assembly pathways (Fig. 6H). Intriguingly, cell adhesion (Kirov et al., 2009; Liu et al., 2018), synapse function (Fromer et al., 2014), cell migration (Senturk et al., 2011; Ishii et al.,
2016) and gliogenesis (Windrem et al., 2017)) have all been highlighted in schizophrenia studies. Therefore, rs3617 may participate in schizophrenia pathogenesis via modulating these pathways.

We also performed KEGG analysis and found that the differentially expressed genes were significantly enriched in MAPK, ECM-receptor interaction, adhesion and PI3K-AKT signaling pathways (Fig. 6I). Intriguingly, the ITIH3 gene encodes the heavy chain submit of the pre-alpha-trypsin inhibitor complex, which binds hyaluronic acid and thereby stabilizes extracellular matrix. Therefore, our results and previous studies provide convergent evidence that alterations in extracellular matrix function might mediate the biological impact of rs3617 in schizophrenia pathogenesis. Additionally, previous studies have revealed pivotal roles of MAPK and PI3K-AKT signaling pathways in schizophrenia (Emamian et al., 2004; Funk et al., 2012). It is thus highly likely that rs3617 may contribute to schizophrenia through regulating these signaling pathways. Based on the above findings, we proposed a possible hypothesis to elucidate the potential role of rs3617 in schizophrenia pathogenesis (i.e., rs3617 may confer schizophrenia risk through affecting ITIH3 protein level and neurodevelopment) (Fig. 7).

3. Discussion

GWASs of schizophrenia have identified over 180 risk loci showing strong associations with schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Li et al., 2017; Pardinas et al., 2018). However, the causal variants and genes remain largely unknown for most of the reported risk loci. Considering that most of the identified risk variants were located in non-coding regions, it is likely that a large proportion of risk variants exert their effects on schizophrenia through modulating gene expression (rather than affecting protein
function). Accordingly, recent studies have mainly focused on non-coding risk variants (Luo et al., 2015; Yang et al., 2018), leaving the coding variants, which likely participate in an illness by altering protein structure and function, relatively less investigated in schizophrenia (Li et al., 2016). In this study, by exploring the associations between the potential causal coding variants and schizophrenia in a large Han Chinese sample, we showed that a non-synonymous SNP in ITIH3 gene, rs3617, was significantly associated with the disease, and therefore was likely an authentic risk variant for schizophrenia.

The successful replication of rs3617 in Han Chinese population provides further evidence that supports its involvement in schizophrenia. Interestingly, rs3617 has the same risk allele (i.e., C allele) in European and Chinese populations, indicating that it is a common risk variant in different ethnic populations. We have also explored the impact of different alleles of rs3617 on ITIH3 function. We show that rs3617 is located in an evolutionarily highly conserved region, and its schizophrenia risk C allele is associated with lower ITIH3 protein expression. Intriguingly, mNSCs stably transfected with different alleles of rs3617 exhibited significant differences in proliferation, differentiation and migration, suggesting that rs3617 may exert its biological effect through modulating neurodevelopment. Transcriptome analysis using mNSCs stably transfected with different alleles of rs3617 has identified potential biological processes (or pathways) affected by rs3617. Specifically, genes differentially expressed between cells with different alleles at rs3617 are enriched in cell adhesion, synapse organization and cell migration processes, confirming the potential impact of rs3617 on neural migration and neurodevelopment. These results provide convergent evidence that rs3617 is likely a causal variant that confers schizophrenia risk through affecting neurodevelopment.
Though PGC predicted that rs3617 might be a potential causal variant, to date the functional consequences and biological effects of rs3617 remain elusive. In this study, we provide convergent lines of evidence that support rs3617 as a causal variant for schizophrenia. First, rs3617 locates in the coding region of ITIH3, and does cause altered amino acid sequence of ITIH3. Second, rs3617 is also associated with schizophrenia in Han Chinese population, further supporting its putative role in schizophrenia. Third, the genomic region harboring rs3617 is highly evolutionarily-conserved, suggesting the functional conservation of this genomic region. Fourth, the allelic differences at rs3617 lead to amino acid change from lysine (corresponding to A allele) to glutamine (corresponding to C allele). Considering that lysine and glutamine have different physiochemical properties (for example, lysine is a charged amino acid while glutamine is uncharged), the differences of amino acid at rs3617 may affect ITIH3 function. Fifth, we show that the C allele of rs3617 is associated with lower ITIH3 protein expression. Sixth, mNSCs stably transfected with different alleles of rs3617 exhibit significant differences in proliferation, migration and differentiation, confirming the functional consequences of rs3617. Seventh, the differentially expressed genes identified in mNSCs stably transfected with different alleles of rs3617 are significantly enriched in MAPK, cell adhesion and PI3K-ATK signaling pathways. Considering the pivotal role of these signaling pathways in schizophrenia (Emamian et al., 2004; Funk et al., 2012), it is possible that rs3617 confers schizophrenia risk through modulating these signaling pathways. Taken together, these convergent lines of evidence suggest that rs3617 is a causal variant for schizophrenia.

Accumulating evidence supports that ITIH3 is likely an authentic risk gene for psychiatric disorders. Genetic variants near ITIH3 have been reported to be associated with schizophrenia.
(Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium, 2011; Yang et al., 2019), bipolar disorder (Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011) and depression (Miyake et al., 2018). In addition to the significant associations observed in European populations, Li et al. also found that a SNP (rs2239547) near the ITIH3/4 locus was significantly associated with schizophrenia in Han Chinese population (Li et al., 2015). In line with the genetic studies, a recent study showed that functional deficiency of Itih3 in mice led to abnormal behaviors, including increased anxiety-like behavior, reduced exploratory activity, altered acoustic startle responses and etc. (Goulding et al., 2019). These behavioral results provide further support for the involvement of ITIH3 in psychiatric disorders.

ITIH3 is a member of the inter-alpha-trypsin inhibitor (ITI) family. ITIH3 encodes inter-alpha-trypsin inhibitor heavy chain H3, a subunit of the pre-alpha-trypsin inhibitor complex (Diarra-Mehrpour et al., 1998). Previous studies have shown that the pre-alpha-trypsin inhibitor complex is involved in extracellular matrix stabilization (Bost et al., 1998; Diarra-Mehrpour et al., 1998). The extracellular matrix has important functions in the developing brain. It is involved in differentiation of NSCs, neuronal migration, the formation of axonal tracts, and the maturation of synapses in the brain (Barros et al., 2010; Franco and Muller, 2011). Considering the important roles of extracellular matrix in brain development and function, it is possible that ITIH3 may confer schizophrenia risk through affecting the function of extracellular matrix, which in turn influences neurodevelopment. Consistent with this, we found that mNSCs stably overexpressing different alleles of rs3617 exhibited significant differences in proliferation, migration and differentiation. Our comparative analysis of the transcriptomes in mNSCs stably overexpressing different alleles at rs3617 also supports the hypothesis that rs3617 may participate in
schizophrenia pathogenesis through affecting neurodevelopment. Intriguingly, this contention is further strengthened by the observation that differentially expressed genes in cells carrying different alleles at rs3617 were significantly enriched in neurodevelopment-related pathways, including cell adhesion, cell migration, gliogenesis and synapse organization. In addition, KEGG analysis highlighted enrichment of these differentially expressed genes in MAPK and PI3K-AKT signaling pathways, two pathways with pivotal roles in schizophrenia (Emamian et al., 2004; Funk et al., 2012).

Of note, we found that the human ITIH3 protein harboring the rs3617-C allele (risk allele) increased the proliferation of mNSCs and promoted the differentiation of mNSCs into MAP2+ and TUJ1+ neuronal cells. Though previous studies have shown that down-regulation of several schizophrenia risk genes (including DISC1 and RELN) inhibited proliferation of NSCs (Won et al., 2006; Mao et al., 2009; Ishizuka et al., 2011), several studies also showed that down-regulation of schizophrenia risk genes (GLT8D1 and VRK2) could increase proliferation of NSCs (Yu et al., 2017; Yang et al., 2018). In the current study, the schizophrenia risk C allele of rs3617 is associated with lower ITIH3 protein expression, and it is possible that human ITIH3 protein harboring C allele at rs3617 increased the proliferation of mNSCs through mechanisms similar to GLT8D1 and VRK2. In addition, we also noticed that migration of mNSCs carrying rs3617-C allele was significantly impaired compared with those carrying rs3617-A allele (Fig. 4). Therefore, the human ITIH3 protein corresponding to the schizophrenia risk C allele of rs3617 likely affects both proliferation and migration processes that are essential for proper neurodevelopment. Finally, we found that differentiation of mNSCs into glial cells was significantly impaired when the cells stably overexpressed ITIH3 and rs3617-C allele (Fig. 5A and B), suggesting that rs3617 may
affect the appropriate ratios of neuronal and glial cells in the brain. Taken together, these data strongly suggest that rs3617 affects neurodevelopment through regulating proliferation, migration and differentiation of NSCs, and thereby contributes to schizophrenia pathogenesis.

In both GTEx and LIBD brain eQTL datasets, no significant association between rs3617 and ITIH3 mRNA expression was observed. A possible explanation is that rs3617 may exert greater effect on the protein levels of ITIH3 while less impact on the mRNA expression. We compared the effect size of different alleles of rs3617 on mRNA and protein levels of ITIH3 in triplicated samples. In HEK293 cells, we found that the effect of rs3617 on ITIH3 mRNA expression was more significant than that on ITIH3 protein. Compared with cells transfected with A allele of rs3617, cells transfected with C allele exhibited a 42% reduction of ITIH3 mRNA, while only 20% reduction of the protein. However, in mNSCs, we found that the effect of rs3617 on ITIH3 protein expression was more significant than that on ITIH3 mRNA expression level. Specifically, the ITIH3 mRNA level was reduced by 25%, and the ITIH3 protein level was reduced by 31% in NSCs transfected with C allele of rs3617 compared with those transfected with A allele. However, it is also possible that rs3617 may exert its main effects through regulating expression of other genes (or affecting ITIH3 activity). More work is needed to examine the potential impact of rs3617 on other genes.

Despite the enlightening discoveries, there are several limitations of this study. First, eight coding SNPs were successfully genotyped in our study, however, only rs3617 showed a genome-wide significant association with schizophrenia. Nevertheless, we noticed that 5 out of the other 7 (including rs950169, rs4584886, rs2288920, rs20551 and rs133335) coding SNPs have the same allelic direction of their effects in both PGC and our sample. Considering that the sample
size in PGC is much larger than the current study, these SNPs may show significant associations with schizophrenia in Chinese population as the sample size increases. In addition, considering the genetic heterogeneity (i.e., different genetic variants in a gene were associated with schizophrenia in different populations) of schizophrenia (Li et al., 2011; Yue et al., 2011), it is also possible that other genetic variants, which were not highlighted in PGC, confer causative risk of schizophrenia in Chinese individuals. More work is thus needed to investigate the association between schizophrenia and other SNPs of these loci in Han Chinese population. Second, we have shown that rs3617 is a functional non-synonymous SNP. However, we still do not know how different amino acids encoded by rs3617 affect the function of ITIH3 protein. It is possible that p.K315Q may influence the biological function of ITIH3 rather than its localization. More work is needed to investigate the precise mechanisms. Third, we investigated the potential functional consequences and biological effects of rs3617 through overexpressing the ITIH3 proteins containing different amino acids at rs3617. Though our results suggest that endogenous ITIH3 expression may not affect our results significantly, potential bias might still be introduced as the use of overexpression vectors leads to much higher levels of ITIH3 protein in these cells than the physiological level. Ideally, additional analyses using genomic editing techniques (e.g., CRISPR-Cas9 mediated single base editing) to obtain cells having different alleles of rs3617 while expressing ITIH3 at physiological level are needed. Fourth, most of the results of this study were from in vitro assays, in vivo validation of the functional consequences of rs3617 using mice carrying different alleles of rs3617 is needed. For example, CRISPR-Cas9 mediated single base editing might allow the generation of these mice. Physiological and pathological evaluation of these mice will then provide pivotal information about the role of rs3617 in schizophrenia pathogenesis. Finally, in the
differentiation assays, although mNSCs underwent differentiation for 4.5 days in differentiation medium, it might still be possible that a very small proportion of progenitor cells (GFAP positive) remained in the final cell populations, which affected the ratio of GFAP positive cells we calculated. Further validation is needed to differentiate these GFAP positive progenitors and glial cells in the final cell populations.

In summary, we have shown that rs3617 is significantly associated with schizophrenia in Han Chinese population (with the same risk allele as in European population), indicating that rs3617 is a common risk variant for schizophrenia across different ethnicities. We have also confirmed that rs3617 is a functional missense SNP, which exerts biological effect on schizophrenia through affecting neurodevelopment. Finally, we have identified genes differentially expressed in mNSCs stably transfected with different alleles of rs3617, and have found significant enrichment of these genes in multiple signaling pathways essential for physiological homeostasis and neurodevelopment (e.g., MAPK and PI3K-AKT signaling pathways), suggesting that rs3617 may confer risk of schizophrenia by modulating these pathways. Our study demonstrates how a coding SNP (rs3617) confers risk of schizophrenia through affecting ITIH3 function, shedding new light on the genetic mechanisms of schizophrenia.

4. Materials and methods

4.1. Study subjects

A total of 4022 unrelated schizophrenia cases and 9270 controls were included in this study. Detailed information about the cases and controls have been described in our previous studies.
Briefly, the samples included in this study were mainly recruited from Hunan (Ma et al., 2013; Zhang et al., 2014) (Institute of Mental Health, National Clinical Research Center for Mental Health Disorders and National Technology Institute of Psychiatry, The Second Xiangya Hospital, Central South University) and Yunnan (Luo et al., 2008; Li et al., 2011) (Yunnan Mental Health Center) provinces (two geographically dispersed regions of China) of China. To guarantee the quality of each subject included in this study, the detailed information (including the onset and course of psychosis, symptoms, family history of psychiatric illnesses, personal and family history report, and medical reports were collected) of each case was collected and carefully assessed by experienced psychiatrists. All the unrelated subjects were of Han Chinese origin and the cases were diagnosed with schizophrenia according to DSM-IV or ICD-10 criteria. Diagnoses were performed by experienced psychiatrists, and individuals with a history of neurological diseases, drug abuse and alcohol abuse were excluded in this study. The ages of cases and controls were 28 ± 13.28 and 38 ± 12.06 years, respectively. Written informed consents were obtained from all participants, and this study was approved by the internal review board of Kunming Institute of Zoology, Chinese Academy of Sciences.

4.2. Genomic DNA extraction and genotyping

We used the standard phenol-chloroform method to extract genomic DNA (from peripheral blood). The 8 potential causal coding SNPs are showed in Table 1. Though 10 potential causal coding SNPs were identified by PGC (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), only 8 non-synonymous exonic SNPs were included in this study, as
rs13107325 is monomorphic in Chinese according to the 1000 Genomes project (Abecasis et al., 2010), and rs2955365 was failed to be genotyped. Genotyping was performed using the SNaPshot approach, which utilizes a single-base extension principle to genotype the desired SNPs. Detailed procedures for DNA extraction and genotyping can be found in our previous study (Luo et al., 2008). In brief, we first amplified the genomic sequence containing the tested SNP by PCR (the PCR primers were provided in Table S6). We then treated the PCR products with Shrimp Alkaline Phosphatase (SAP) and Exonuclease I (ExoI) to remove the unincorporated dNTPs and primers. PCR was conducted again using the purified products, SNaPshot multiplex mix, genotyping primers (the genotyping PCR primers were provided in Table S7) and ddNTPs. The 3′ end of the genotyping primer stops just one base prior to the test SNP, so the polymerase extends the primer by one nucleotide (adding a single ddNTP to its 3′end). As the ddNTPs were labeled with fluorescence, the fluorescence color readout reports which base was added. Electrophoresis was performed with 3730 DNA analyzer (Applied Biosystems, Japan) and the data were analyzed with GeneMapper software. Genotypes of each SNP were manually checked. The genotyping results were validated with Sanger sequencing through sequencing 100 randomly selected individuals and no genotyping errors were observed.

4.3. Genetic association analysis

The Hardy-Weinberg equilibrium (HWE) for each SNP was assessed with PLINK (Purcell et al., 2007). Allelic association test of each SNP was conducted with Chi-square test (implemented in PLINK) (Purcell et al., 2007). Meta-analysis was performed as previously described (Li et al., 2018) and the fixed-effect model was used.
4.4. Sequence conservation and 3D structure analyses

The protein sequences of 35 vertebrate species surrounding p.K315Q (rs3617) were extracted from the UCSC genome browser (https://genome.ucsc.edu/). The 3D structures of ITIH3 containing lysine and glutamine amino acids at rs3617 were simulated with SWISS-MODEL (https://swissmodel.expasy.org/).

4.5. Linkage disequilibrium (LD) analysis

We examined the LD patterns of the genomic regions containing the 8 tested non-synonymous SNPs using the genotype data (phase 3) from the 1000 Genomes project (The 1000 Genomes Project Consortium, 2015). Briefly, genotype data (±5kb of the non-synonymous SNP) of 99 Europeans (CEU) and 103 Chinese (CHB) subjects were downloaded. LD maps were generated with Haploview (Barrett, 2009) and LD values were presented as $r^2$. We performed SNP filtering before plotting LD plot. Our SNP filtering parameters are as follows: Hardy-Weinberg P-value ≥ 0.01, genotyping rate ≥ 75% and minor allele frequency ≥ 0.01. We used the default algorithm (based on confidence interval, which was developed by Gabriel et al. (Gabriel et al., 2002)) implemented in Haplovview software (Barrett et al., 2005; Barrett, 2009) to define LD blocks. This algorithm first labels comparison between variants as “strong LD”, “inconclusive” and “strong recombination” based on 95% confidence bounds on D prime ($D'$) for each comparison. Then a block is defined if 95% of informative comparisons are labeled as “strong LD”. More detailed information about the definition of LD blocks can be found in previous papers (Gabriel et al., 2002; Barrett et al., 2005). To make sure that the generated LD plot centered on the exonic variant,
we firstly extracted the SNP surrounding the exonic variant (± 10 kb). We then selected SNPs manually to guarantee that each exonic SNP has the same (or similar) number of upstream and downstream variants.

4.6. Expression quantitative trait loci (eQTL) analysis

To explore if rs3617 was associated with gene expression in human brain tissues, we examined the associations between rs3617 and gene expression using the brain eQTL data from GTEx (The GTEx Consortium, 2015) and LIBD (eQTL of dorsolateral prefrontal cortex tissues from 412 subjects) (Jaffe et al., 2018) datasets. Detailed information about the GTEx and LIBD brain eQTL data can be found in the original publications (The GTEx Consortium, 2015; Jaffe et al., 2018).

4.7. Construction of ITIH3 expression vector and transfection

To explore the impact of different alleles of rs3617 on ITIH3 protein expression, we constructed ITIH3 expression constructs through cloning the full-length coding sequence of human ITIH3 (2670 bp) into the pEGFP-N2 vector. Two expression vectors carrying different alleles (either C or A allele) at rs3617 and the same sequence otherwise were constructed. The sequence of ITIH3 was inserted into XhoI and KpnI restriction sites of pEGFP-N2, and the inserted sequence of ITIH3 was validated by direct sequencing. To quantify ITIH3 expression, GFP was fused to the C-terminus of ITIH3. Equal amount of ITIH3 expression vectors (containing either C or A allele of rs3617) were transfected into HEK293 cells using the polyethylenimine (PEI) (Polyscience, USA) reagent.
4.8. Isolation and characterization of mouse neural stem cells

Mouse neural stem cells (mNSCs) were isolated as previously described (Azari et al., 2011; Yang et al., 2018). Briefly, brains of mouse embryos (E13.5) were harvested and ganglionic eminences were dissected. The dissected brains were micro-dissected and brain tissues were dissociated thoroughly through repeated pipetting. The cell suspension was then transferred to a new tube and centrifuged for 110 g for 5 mins at room temperature. The supernatant was discarded and the cells were re-suspended to make a homogeneous single cell suspension. The cells were then plated at the density of ~3×10^5 cells/mL in NSC medium (STEMCELL, Canada). The mNSCs were cultured in NeuroCult™ Basal Medium (STEMCELL), which was supplemented with 10% proliferation supplement (STEMCELL, NeuroCult™ Proliferation Supplement (Mouse & Rat)), 20 ng/mL of epidermal growth factor (EGF, STEMCELL), 20 ng/mL of basic fibroblast growth factor (bFGF, STEMCELL) and 2 μg/mL heparin (STEMCELL). After culturing for 3-5 days, neurospheres formed by mNSCs were used for further experiments and passaging. We used three well-characterized NSCs markers (SOX2, PAX6 and NESTIN) to characterize the isolated mNSCs.

4.9. Construction of NSCs stably overexpressing ITIH3

To investigate if the different alleles of rs3617 have different effects on proliferation, differentiation and migration of mNSCs, we generated mNSCs that overexpress ITIH3 stably. Briefly, the full-length coding sequence of human ITIH3 (2673 bp, containing TGA terminator) was cloned into pCDH lentiviral vector at multiple clone sites EcoRI and NotI. Flag-tag (3×) was fused to the N-terminus of ITIH3. To produce functional lentiviral particles, the cloned pCDH
vectors (containing either A or C allele of rs3617), psPAX2 and pMD2G vectors (obtained from
the lab of Prof. Yongbin Chen) were co-transfected into HEK293T cells using PEI. Cell culture
supernatants were collected at 48 h and 72 h after transfection respectively. The supernatants were
then concentrated and used to infect mNSCs. 24 h after infection, mNSCs were treated with
puromycin (2 µg/mL) for 14 days to select cell that were stably transfected with ITIH3
overexpression vectors.

4.10. Cell culture

The HEK293 and HEK293T cells, originally from the American Type Culture Collection (ATCC),
were obtained from Shanghai Cell Bank, Chinese Academy of Sciences. Cells were cultured in
Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum (FBS,
Gibco, USA) and penicillin/streptomycin (100 U/mL penicillin and 0.1 mg/mL streptomycin,
HyClone, USA) at 37°C with 5% CO₂. Mycoplasma was tested periodically to ensure that no
mycoplasma contamination was detected in these cell lines during the current study.

4.11. BrdU proliferation assays

To compare the effect of different alleles of rs3617 on proliferation, we performed
5-Bromo-2-deoxyUridine (BrdU, a nucleoside analogue) incorporation assays. The HEK293 cells
were plated in 24-well plates at a density of 4×10⁴ per well. After 24 h, equal amount of ITIH3
overexpression vectors with either C allele or A allele of rs3617 were transfected into the HEK293
cells. When the cell density reached ~80-90% confluence (48 h after transfection), 10 µM BrdU
(Sigma, Japan) was added to each well and incubated for 1 h. The cells were then fixed with 4%
PFA and permeabilized with 0.3% Triton X-100. After three washes with PBS, 2M HCl was added to denature DNA at 37°C for 30 min. Cells were then blocked with blocking buffer at room temperature for 1 h, followed by incubation with the primary antibody (rat anti-BrdU, Novas, USA, 1:1000, Cat# NB500-169) overnight at 4°C and secondary antibody (Life Technology, USA, 1:500, Cat# A10522) for 1 h at room temperature. We also carried out BrdU incorporation assays for mNSCs stably overexpressing ITIH3. Immunofluorescence staining images were obtained with confocal microscope (Olympus, Japan, FV1000). The number of cells were counted using the ImageJ software (https://imagej.nih.gov/ij/). A total of 10 immunostaining images from three biological replicates were used for cell counting.

4.12. CCK-8 assays

The HEK293 cells were seeded into a 96-well plates (3×10^3 per well). After 24 h, cells were transfected with the following expression vectors: pEGFP, pEGFP-ITIH3-C allele, pEGFP-ITIH3-A allele. These vectors express GFP, ITIH3 protein with Glutamine (p.315Q) at rs3617, ITIH3 protein with Lysine (p.315K) at rs3617, respectively. 48 h after transfection, Cell Counting Kit-8 (CCK-8, Beyotime, China) was used to measure the proliferation rate of cells following the manufacturer’s instruction. Spectrophotometer was used to measure the absorbance of each well at wavelengths of 450 nm and 630 nm. The cell viability was quantified by subtracting the background absorbance at 630 nm from the absorbance at 450 nm. We also tested the impact of different alleles of rs3617 on the proliferation of mNSCs using the same method. Five biological replicates were used for each experimental condition.
4.13. Differentiation of NSCs into neurons and glia cells

The mNSCs stably overexpressing ITIH3 and carrying either C or A allele at rs3617 were seeded (1×10^5 cells/well) into 24-well plates pre-coated with 10 µg/mL laminin. To differentiate the mNSCs into neuronal and glial cells, we used the following differentiation medium: DMEM/F12 (Gibco), B27 (Gibco), N-2 (Gibco) and heparin (STEMCELL Technologies). After differentiation for 4.5 days, immunostaining was performed to count the number of neuronal and glial cells. Briefly, cells were fixed with 4% PFA and permeabilized with PBS containing 0.3% Triton-X100, then blocked with blocking buffer (Beyotime) for 1 h at room temperature. The cells were then incubated with primary antibodies overnight at 4°C, followed with incubation with species-specific fluorescent secondary antibodies at room temperature for 1 h. The primary antibody used in differentiation assays were: GFAP (Sigma, 1:2000, Cat# G9269), TUJ1 (Abcom, England, 1:500, Cat# ab78078) and MAP2 (Millipore, USA, 1:500, Cat# AB5622). The secondary antibodies were: Alexa Fluor 488 donkey anti-rabbit (Life Technology, 1:500), Alexa Fluor 568 donkey anti-mouse (Life Technology, 1:500). A total of 11 immunostaining images from three biological replicates were used to count GFAP positive cells. A total of 15 immunostaining images from three biological replicates were used to count MAP2 and TUJ1 positive cells, respectively.

4.14. Migration assays of NSCs

Migration assay was performed as previously described (Kong et al., 2008; Li et al., 2017). Briefly, neurospheres with similar diameters were selected and plated in 24-well plates (pre-coated with laminin). Only one neurosphere was seeded into each well. After culturing neurospheres in NeuroCult™ Basal Medium (STEMCELL) supplemented with 10% differentiation supplement...
(STEMCELL, NeuroCult™ Differentiation Supplement (mouse)) for 24 h. The migration distance of neurosphere was quantified through subtracting the radius of the original neurosphere from the radius of the neurosphere at 24 h using the ImageJ software. Migration distance of five neurospheres were analyzed for each group (i.e., NSCs stably transfected with the control vector pCDH, ITIH3 expressing vectors containing C allele at rs3617, and ITIH3 expressing vector containing A allele at rs3617).

4.15. Western blotting

Total proteins were extracted using the RIPA buffer (Thermo, USA) containing 1% PMSF (Beyotime). Protein concentration was quantified using the bicinchoninic acid (BCA) Protein Assay Kit (Thermo). After running on 10% SDS-polyacrylamide gel electrophoresis, proteins were transferred to a polyvinylidene difluoride (PVDF) membrane. 5% defatted milk was used to block the nonspecific binding. The PVDF membrane was then incubated with primary antibodies overnight at 4°, followed by three washes with TBST, then were incubated with secondary antibodies for 1 h at room temperature. The SuperSignal West Pico Chemiluminescent Substrate (Thermo) was used for protein detection. The primary antibodies used in the Western blotting assay included: mouse anti-GFP (Santa Cruz, USA, 1:100, Cat# sc-9996), Rabbit anti-ACTB (Sangon, China, 1:1000, Cat# D110001-0200), Rabbit anti-FLAG (Sigma, 1:1000, Cat# F7425). The secondary antibodies were: HRP-labeled Goat Anti-Mouse IgG (H+L) (Beyotime, 1:1000, Cat# A0216), HRP-labeled Goat Anti-Rabbit IgG (H+L) (Beyotime, 1:1000, Cat# A0208).

4.16. Protein degradation assays (Cycloheximide chase analysis)
To test if lysine and glutamine corresponding to different alleles of rs3617 at the 315th amino acid of ITIH3 affected its stability, we conducted cycloheximide (CHX) chase analysis. CHX is a translation inhibitor that inhibits eukaryotic translation and protein synthesis. As protein translation is inhibited by CHX, no new proteins can be synthesized after CHX treatment. The CHX Chase assay is therefore widely used to analyze protein degradation (Buchanan et al., 2016).

Briefly, 1.5×10⁵ HEK293T cells were plated in each well of 12-well plates. The day after cell plating, cells were transfected with either ITIH3-rs3617-A or ITIH3-rs3617-C expression plasmids. Glutathione-S-Transferase (GST) plasmids were co-transfected to control for transfection efficiency. 24 h after transfection, cells were treated with 50 μg/mL CHX at indicated time points and harvested for Western blotting analysis.

### 4.17. Real-time quantitative PCR

Total RNA was extracted with Trizol Reagent (life technologies). After removing potential genomic DNA contamination using the PrimeScript™ RT reagent Kit with gDNA Eraser (Perfect Real Time) (Takara, Japan), reverse transcription was performed using Takara Reverse Transcription System (PrimeScript™ RT reagent Kit). Real-time quantitative PCR (qPCR) was performed using TB Green Premix Ex Taq II kit (Takara). Beta-actin (ACTB) was used as internal control. Gene expression levels were analyzed and presented using the $2^{-ΔΔCt}$ method (Livak and Schmittgen, 2001). The Real-time qPCR primers were provided in Table S8.

### 4.18. Transcriptome analysis with RNA sequencing

Total RNA of mNSCs stably overexpressing ITIH3 (containing either C or A allele at rs3617) was
extracted using Trizol reagent (Invitrogen, USA). RNA concentration was quantified using Qubit®3.0 Flurometer (Life Technologies, CA, USA). RNA purity was determined using the NanoPhotometer® Spectrophotometer (IMPLEN, CA, USA), and RNA integrity was assessed with the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). 3 µg total RNA from each sample was used to construct sequencing libraries according to the recommendations of NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (#E7530L, NEB, USA), and different index codes were used to generate sequencing libraries. After initial quantification, Oligo (dT) and fragmentation buffer were used to enrich and then fragment the mRNA. The fragmented mRNA was then reversed transcribed into cDNA using random primers. The purified double strand cDNA was then end repaired, tailed and ligated. Paired-end RNA sequencing (RNA-Seq) (Illumina, USA) was performed to quantify gene expression levels.

Clean reads were mapped to the mouse genome (GRCm38) using hisat2 (Kim et al., 2015) (v2.1.0) with default parameters. Stringtie (Pertea et al., 2015) (v1.3.4) was used to assemble transcripts with default parameters. DEseq2 (Love et al., 2014) (implemented in R (v3.5.1)) was used to identify differentially expressed genes. Bonferroni correction was conducted to correct P values. Genes with |log2FC| ≥ 1 and Padj ≤ 0.05 were defined as significantly differentially expressed genes. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed using ClusterProfiler (Yu et al., 2012) in R (v3.5.1).

4.19. **Dosage effect assay**

Equal amount HEK293 cells (1×10^5) were plated in 96-well plate and cultured for 18 hours. Then the cells were transfected with different amounts of ITIH3 overexpressing vector (0.5, 1.0, 1.5 and 2.0 µg, respectively). Forty-eight hours post transfection, cell viability was determined using the
Cell Counting Kit-8 (CCK-8) assays.

### 4.20. Statistical analysis

The associations between the genotyped SNPs and schizophrenia were conducted with PLINK (Chi-square test) (Purcell et al., 2007). Meta-analysis was performed as previously described (Li et al., 2018) and the fixed-effect model was used. For ITIH3 protein quantification experiments, proliferation, migration and differentiation of NSCs assays, and qPCR experiments, two-tailed Student's t test was used and significance threshold was set at $P < 0.05$. 
Acknowledgements

This study was equally supported by the National Nature Science Foundation of China (31970561 and 31722029 to X.J.L) and the National Key Research and Development Program of China (Stem Cell and Translational Research) (2016YFA0100900), the Innovative Research Team of Science and Technology department of Yunnan Province (2019HC004) and the Key Research Project of Yunnan Province (2017FA008 to X.J.L).

Supplementary data

Supplementary data to this article can be found online version.
References


Pardinas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S. E., Bishop, S., Cameron, D., Hamshere, M. L., Han, J., Hubbard, L., Lynham, A., Mantripragada,


Saha, S., Chant, D. and McGrath, J. 2007. A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time? Arch Gen Psychiatry 64, 1123-1131.


Figure legends

**Fig. 1.** rs3617 is located in an evolutionary highly conserved region. A: Gene structure of ITIH3 and the location of rs3617. SNP rs3617 locates in the exon 9 of ITIH3 gene. B: SNP rs3617 locates in the vWFA domain. The C allele (risk allele) of rs3617 encodes glutamine, while the A allele encodes lysine. C: SNP rs3617 is located in an evolutionarily highly conserved region. The amino acid at rs3617 site is lysine in most species. However, a new amino acid (glutamine) has emerged in humans. D and E: 3D structure prediction of ITIH3 protein using the SWISS-MODEL website. The locations of amino acids (lysine and glutamine) encoded by different alleles of rs3617 are showed (arrowheads).

**Fig. 2.** The C allele of rs3617 is associated with lower ITIH3 protein expression. A: ITIH3 expression vectors (2.5 μg) containing C (pEGFP-ITIH3-C) and A (pEGFP-ITIH3-A allele) allele of rs3617 were transiently transfected into HEK293 cells. Western blotting (WB) was used to quantify ITIH3 protein expression and Actin was used as an internal control. The normalized ITIH3 expression values (GFP/Actin) were compared to test if different alleles of rs3617 affected ITIH3 expression. Total proteins were collected at three time points (24, 48 and 72 h) after transfection, and WB was performed. B–D: Quantification results of the ITIH3 expression at 24 h (B), 48 h (C) and 72 h (D) post transfection. Data are represented as Mean ± SD, and P values were calculated by two-tailed Student's t test, n = 6. E: ITIH3 expression vectors containing different alleles of rs3617 were constructed and Flag was fused to the N terminus of the ITIH3 coding sequence. The constructed expression vectors (2.5 μg) were transfected into HEK293 cells. 48 h after transfection, total protein was extracted and ITIH3 expression was quantified. Western
blotting showed that the C allele of rs3617 was associated with lower ITIH3 protein expression. F: Quantification of the WB results showed in (E). Data are represented as mean ± SD, and P values were calculated by two-tailed Student's t test, n = 3. G: Proteins of mNSCs stably overexpressing different alleles of rs3617 were extracted, and WB showed that the C allele of rs3617 was associated with lower ITIH3 protein expression in mNSCs. H: Quantification of the WB results showed in (G), n = 3. *, P < 0.05, ***, P < 0.001.

Fig. 3. The mNSCs stably transfected with different alleles of rs3617 exhibited significant differences in proliferation. A–E: Isolation and characterization of the mNSCs. Immunofluorescence staining with three well-characterized NSC markers (PAX6, NESTIN and SOX2) confirmed that the cells isolated were NSCs. F: BrdU incorporation assay showed that NSCs stably transfected with rs3617-C allele had significantly higher proliferation rates compared with those transfected with rs3617-A allele. G: The quantified results of the BrdU incorporation assay. H: Consistent with the BrdU assay, CCK-8 assay also revealed that NSCs stably transfected with C allele had higher proliferation rates. I: showed HEK293 cells transiently transfected with C allele have significantly higher proliferation rates compared with cells transfected with the A allele. J: Quantified results of the BrdU incorporation assay in I. K: CCK-8 assay revealed that HEK293 cells transiently transfected with ITIH3 expression vector containing C allele of rs3617 had higher proliferation rates compared with cells containing the A allele. Data are represented as mean ± SD, and P values are calculated by two-tailed Student's t test, n = 3; a total of 10 randomly selected
immunostaining images were used for cell counting for G, n = 5 for H, n=3 and a total of 6
immunostaining images were used for cell counting for J, n = 5 for K. *, P < 0.05, **, P < 0.01,
***, P < 0.001.

Fig. 4. NSCs stably transfected with different alleles of rs3617 exhibited significant
differences in migration. A: Representative photographs for the migration assay. Neurospheres
with similar diameters were selected and seeded into laminin-coated 24-well plates. After
culturing neurospheres in differentiation medium for 24 h, the migration distances of neurospheres
were measured by subtracting the radius of the selected neurosphere at 24 h from the radius of the
original neurosphere. B: Quantified results of the migration assay in A. Data are represented as
Mean ± SD, and P value was calculated by two-tailed Student’s t test, 5 independent neurospheres
were analyzed for each group. Scale bar represents 500 μm in A. *, P < 0.05.

Fig. 5. The mNSCs stably transfected with different alleles of rs3617 exhibited significant
differences in differentiation. A: Differentiation of mNSCs stably transfected either C or A allele
of rs3617 into astrocytes. Representative immunostainings for GFAP, a marker for astrocytes. B:
The quantification of results in A. Compared with mNSCs stably transfected with A allele of
rs3716, the differentiation of cells stably transfected with rs3617-C allele into astrocytes (GFAP
positive) was significantly impaired (B). Differentiation of mNSCs into mature neurons. C and D:
Representative immunostaining images (C) and quantified results (D) of MAP2 (a marker for mature neurons) positive cells. Compared with mNSCs stably transfected with A allele of rs3716, the differentiation of mNSCs stably expressing rs3617-C allele into astrocytes (GFAP positive) was significantly increased. E and F: Representative immunostaining images (E) and corresponding quantification (F) of immature postmitotic neurons which are TUJ1 positive. Data are represented as mean ± SD, and P values were calculated by two-tailed Student’s t test, \( n = 3 \) and a total of 11 immunostaining images were used for cell counting for B, \( n = 3 \) and a total of 15 immunostaining images were used for cell counting for D, \( n = 3 \) and a total of 15 immunostaining images were used for cell counting for F. Scale bar represents 50 μm in A, C and E. **, \( P < 0.01 \), ***, \( P < 0.001 \).

**Fig. 6. Transcriptome analysis revealed differentially expressed genes between mNSCs stably transfected with C and A allele of rs3617. A:** Expression heatmap showed the differentially expressed genes (\( n = 669 \)) between mNSCs stably transfected with C and A allele of rs3617. B: The top 30 genes showing the most significant expression differences. C–G: Quantitative PCR validation of RNA-seq results. Genes marked with red color in (B) were selected for qPCR verification. These genes included Peg10 (C), Sgce (D), Dbx2 (E), Rasgrp3 (F) and Syt1 (G). Data are represented as mean ± SD, and P values were calculated by two-tailed Student’s t test. Average results of three independent biological replicates are presented (three technical replicates for each biological replicate). *, \( P < 0.05 \), **, \( P < 0.01 \), ***, \( P < 0.001 \). H: Gene ontology (GO)
analysis for differentially expressed genes identified by RNA-seq. The differentially expressed genes were significantly enriched in neurodevelopment-related pathways, including cell adhesion, synapse organization, cell migration and gliogenesis. I: KEGG analysis for differentially expressed genes identified by RNA-seq. The differentially expressed genes were significantly enriched in MAPK, ECM-receptor interaction, adhesion and PI3K-AKT signaling pathways.

Fig. 7. The working model of rs3617 in schizophrenia pathogenesis. The schizophrenia risk C allele of rs3617 down-regulates ITIH3 expression. Lower ITIH3 expression promotes proliferation of NSCs while inhibits their migration. Lower ITIH3 expression promotes differentiation of NSCs into glial cells while inhibits their differentiation into neurons. SNP rs3617 also alters schizophrenia-associated pathways. These results together suggest that rs3617 may confer risk of schizophrenia through affecting neurodevelopment.
<table>
<thead>
<tr>
<th>SNP</th>
<th>CHR:BP</th>
<th>Gene</th>
<th>A1/A2&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Amino acid Change&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Freq&lt;sub&gt;Cases&lt;/sub&gt;&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Freq&lt;sub&gt;Controls&lt;/sub&gt;&lt;sup&gt;d&lt;/sup&gt;</th>
<th>OR&lt;sup&gt;e&lt;/sup&gt;</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs3176443</td>
<td>1:177247854</td>
<td>FAM5B</td>
<td>G/C</td>
<td>V390L</td>
<td>0.172</td>
<td>0.174</td>
<td>0.987</td>
<td>7.15E-01</td>
</tr>
<tr>
<td>rs3617</td>
<td>3:52833805</td>
<td>ITIH3</td>
<td>A/C</td>
<td>K315Q</td>
<td>0.393</td>
<td>0.446</td>
<td>0.803</td>
<td>8.36E-16</td>
</tr>
<tr>
<td>rs3774729</td>
<td>3:63982082</td>
<td>ATXN7</td>
<td>G/A</td>
<td>M862V</td>
<td>0.502</td>
<td>0.486</td>
<td>1.067</td>
<td>1.48E-02</td>
</tr>
<tr>
<td>rs950169</td>
<td>15:84706461</td>
<td>ADAMTSL3</td>
<td>T/C</td>
<td>I1660T</td>
<td>0.062</td>
<td>0.064</td>
<td>0.978</td>
<td>6.80E-01</td>
</tr>
<tr>
<td>rs4584886</td>
<td>17:17896205</td>
<td>LRR4C8</td>
<td>C/T</td>
<td>R191W</td>
<td>0.126</td>
<td>0.129</td>
<td>0.971</td>
<td>4.68E-01</td>
</tr>
<tr>
<td>rs2288920</td>
<td>19:50091798</td>
<td>PRG2</td>
<td>T/G</td>
<td>C116G</td>
<td>0.198</td>
<td>0.208</td>
<td>0.943</td>
<td>7.82E-02</td>
</tr>
<tr>
<td>rs20551</td>
<td>22:41548008</td>
<td>P300</td>
<td>G/A</td>
<td>V997I</td>
<td>0.080</td>
<td>0.075</td>
<td>1.070</td>
<td>1.74E-01</td>
</tr>
<tr>
<td>rs133335</td>
<td>22:42416056</td>
<td>WBP2NL</td>
<td>A/G</td>
<td>D121G</td>
<td>0.153</td>
<td>0.162</td>
<td>0.935</td>
<td>6.98E-02</td>
</tr>
</tbody>
</table>

<sup>a</sup>A1/A2 represent the minor allele/major allele of each SNP.<br><sup>b</sup>Amino acid change means different alleles encode different amino acids.<br><sup>c</sup>Freq<sub>Cases</sub> is the frequency of A1 allele in cases.<br><sup>d</sup>Freq<sub>Controls</sub> is the frequency of A1 allele in controls.<br><sup>e</sup>OR (odds ratio) was based on A1 allele.
The image shows a comparison of gene expression and BrdU incorporation in m-NSCs (multipotent neural stem cells) under different conditions.

**Legend:**
- DAPI: DNA staining
- PAX6: Proximal axis determinant 6
- NESTIN: Nestin
- SOX2: Sex-determining region Y box 2
- Merge: Combined images
- DAPI/BrDU: Double staining for DAPI and BrdU

**Table F:**
- rs3617.A and rs3617.C conditions are shown side by side.

**Graphs G and H:**
- Graph G: Comparison of BrdU incorporation (as % of DAPI) for rs3617.A and rs3617.C.
- Graph H: Comparison of OD450 values for different conditions.

**Graphs J and K:**
- Graph J: Comparison of BrdU incorporation (as % of DAPI) for different conditions.
- Graph K: Comparison of OD450 values for different conditions.

Scale bars: 100μm for A-E and 50μm for F-I.
<table>
<thead>
<tr>
<th>Time</th>
<th>pCDH</th>
<th>ITIH3-rs3617-A</th>
<th>ITIH3-rs3617-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0h</td>
<td>![Image](72x541 to 529x723)</td>
<td>![Image](72x541 to 529x723)</td>
<td>![Image](72x541 to 529x723)</td>
</tr>
<tr>
<td>24h</td>
<td>![Image](72x541 to 529x723)</td>
<td>![Image](72x541 to 529x723)</td>
<td>![Image](72x541 to 529x723)</td>
</tr>
</tbody>
</table>

**B**

![Graph showing migration distance](72x541 to 529x723)

* Indicates a statistically significant difference.

**Migration Distance (µm)**

- pCDH
- ITIH3-rs3617-A
- ITIH3-rs3617-C
Genetic association study

Schizophrenia cases  
N=4,022

Controls  
N=9,270

rs3617(C/A)

Functional characterization

Risk allele (C)

ITIH3  ITIH3  ITIH3

Non-risk allele (A)

ITIH3  ITIH3  ITIH3

Lower ITIH3 protein

NSCs

Promotes Proliferation
Inhibits Migration

MAPK signaling pathway
PI3K-Akt signaling pathway
Gliogenesis
Cell migration

Altered schizophrenia associated pathways

Abnormal neurodevelopment

Schizophrenia

Inhibits

Astrocytes

Promotes

Neurons