Journal of Psychiatric Research 83 (2016) 168–175

Contents lists available at ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/psychires

Identification of *SLC25A37* as a major depressive disorder risk gene

Yong-Xia Huo [a, 1], Liang Huang [b, 1], Deng-Feng Zhang [a], Yong-Gang Yao [a], Yi-Ru Fang [c], Chen Zhang [c, **], Xiong-Jian Luo [a, *]

[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
[b] First Affiliated Hospital of Gannan Medical University, Ganzhou, Jiangxi 341000, China
[c] Division of Mood Disorders, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China

ARTICLE INFO

Article history:
Received 9 May 2016
Received in revised form
5 September 2016
Accepted 8 September 2016

Keywords:
Major depressive disorder (MDD)
SLC25A37
Genetic association
Expression quantitative trait loci (eQTL)
Integrative analysis
Expression

ABSTRACT

Major depressive disorder (MDD) is one of the most prevalent and disabling mental disorders, but the genetic etiology remains largely unknown. We performed a meta-analysis (14,543 MDD cases and 14,856 controls) through combining the GWAS data from the Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium and the CONVERGE consortium and identified seven SNPs (four of them located in the downstream of *SCL25A37*) that showed suggestive associations ($P < 5.0 \times 10^{-7}$) with MDD. Systematic integration (*Sherlock* integrative analysis) of brain eQTL and GWAS meta-analysis identified *SCL25A37* as a novel MDD risk gene ($P = 2.22 \times 10^{-6}$). A cis SNP (rs6983724, ~28 kb downstream of *SCL25A37*) showed significant association with *SCL25A37* expression ($P = 1.19 \times 10^{-9}$) and suggestive association with MDD ($P = 1.65 \times 10^{-7}$). We validated the significant association between rs6983724 and *SCL25A37* expression in independent expression datasets. Finally, we found that *SCL25A37* is significantly down-regulated in hippocampus and blood of MDD patients ($P = 3.49 \times 10^{-3}$ and $P = 2.66 \times 10^{-13}$, respectively). Our findings implicate that the *SCL25A37* is a MDD susceptibility gene whose expression may influence MDD risk. The consistent down-regulation of *SCL25A37* in MDD patients in three independent samples suggest that *SCL25A37* may be used as a potential biomarker for MDD diagnosis. Further functional characterization of *SCL25A37* may provide a potential target for future therapeutics and diagnostics.

1. Introduction

Major depressive disorder (MDD) is a complex mental disorder characterized by persistent and pervasive low mood, including low self-esteem, loss of interest or pleasure, and feelings of personal worthlessness. The lifetime prevalence of MDD is around 15% (Hasin et al., 2005; Kessler et al., 2003), which makes it one of the most prevalent mental disorders. MDD has a high mortality and significant long-term morbidity (Angst et al., 2002; Lopez et al., 2006). Persons with MDD have a high risk for suicide and approximately 11% MDD patients die from suicide (Wulsin et al., 1999). The economic and social burden of MDD is particularly great (Ferrari et al., 2013; Olchanski et al., 2013). For example, the cost of MDD in USA (including direct costs such as outpatient, inpatient, drugs, and long-term care) and non-health care costs (such as law enforcement, reduced workplace productivity, and unemployment) was estimated to be $173.2 billion in 2005 (Greenberg et al., 2015). However, this number was rapidly increased to $210.5 billion in 2010 (rose by 21%) (Greenberg et al., 2015). With the continuously increase of cost, MDD poses a major global health and economy challenge.

Though MDD affects millions of people and is a leading cause of disability worldwide, the pathophysiology of MDD remains largely unknown. Accumulating evidence indicated that both genetic and environmental factors are involved in the pathogenesis of MDD (CONVERGE consortium*, 2015; Flint and Kendler, 2014; Lesch, 2004; Subbarao et al., 2008; Sullivan et al., 2012). The heritability of MDD was estimated around 0.32 (Lubke et al., 2012; Sullivan

* Corresponding author. Key Laboratory of Animal Models and Human Disease Mechanisms, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650204, China.
** Corresponding author.
E-mail addresses: zhangchen645@gmail.com (C. Zhang), luoxiongjian@mail.kiz.ac.cn (X.-J. Luo).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.jpsychires.2016.09.011
0022-3956/

et al., 2000), indicating that genetic factors play a role in MDD. To elucidate the genetic basis of MDD, numerous genetic studies have been conducted in different human populations (Kohli et al., 2011; Lewis et al., 2010; Muglia et al., 2010; Rietschel et al., 2010; Ripke et al., 2012; Shi et al., 2011; Wray et al., 2011). Nevertheless, only very limited risk genetic variants or susceptibility genes have been identified and validated (CONVERGE consortium*, 2015). The advent of genome-wide association studies (GWAS) provides a chance to decipher the genetic mechanisms of MDD. Due to the dramatic increase in sample size and genotyping throughput, GWAS can identify genetic variants with small effect (CONVERGE consortium*, 2015). In 2012, the Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium (Ripke et al., 2012) conducted a large GWAS analysis of MDD. In the first stage, more than 1.2 million SNPs were analyzed in 9240 MDD cases and 9519 controls. In the second stage, 554 SNPs ($P < 0.001$) from the first stage were replicated in independent samples, including 6783 MDD cases and 50,695 controls. Despite the fact that 76,237 individuals (including 16,023 MDD cases) were included (Ripke et al., 2012), no genome-wide significant SNP was identified, strongly suggest that the genetic architecture of MDD is much complex than we had thought.

Sample size, genetic and phenotypic heterogeneity (there are several subtypes of MDD) may be the possible reasons for the failure of identification of genome-wide significant variants for MDD. To minimize the genetic and phenotypic heterogeneity, the CONVERGE consortium collected a relatively homogenous Chinese sample (5303 MDD cases and 5337 controls) and performed whole-genome sequencing at low coverage (CONVERGE consortium*, 2015). Two genome-wide significant loci (one near the *SIRT1* gene and the other one in an intron of the *LHPP* gene) were identified by the CONVERGE consortium. Though the CONVERGE consortium has successfully identified two MDD risk loci, much of the heritability of MDD remains unknown. More importantly, the identified MDD risk variants reside in non-coding regions, with limited annotation and no obvious functional consequence. Therefore, how these susceptibility variants contribute to MDD risk remains elusive. In addition, many loci may have small effects and fail to reach the genome-wide significance in a single GWAS study. Thus, mining the potential effect of the weak GWAS associations (e.g., variants with $P < 10^{-5}$) may help to uncover the missing heritability of MDD.

Here we carried out a meta-analysis by combining the GWAS data from the Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium (Ripke et al., 2012) and the CONVERGE consortium (CONVERGE consortium*, 2015). We then systematically integrated genetic association signals from the meta-analysis and brain expression quantitative trait loci (eQTL) data through using a Bayesian statistical framework (*Sherlock*), and identified *SLC25A37* as a novel MDD risk gene. To further characterize the potential role of *SLC25A37* in MDD etiology, we attempted to validate our results in independent gene expression datasets, and investigated the expression of *SLC25A37* in MDD patients and healthy controls in independent samples. Our consistent and convergent results suggest that *SLC25A37* may have a role in the etiology of MDD.

2. Materials and methods

2.1. MDD GWAS data and meta-analysis

Sample size is an important factor for GWAS of MDD. New MDD risk variants (or loci) may be identified with the increase of sample size. To this aim, we combined two large-scale GWAS datasets of MDD. The first GWAS dataset was from the Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium (Ripke et al., 2012). In brief, this study represents a mega-analysis of genome-wide association studies for major depressive disorder. The SNP associations (a total of 1,235,109 SNPs) from the discovery phase (including 9240 MDD cases and 9519 controls, all of the subjects were of European ancestry) (hereafter referred as European sample) were used for meta-analysis. The second GWAS dataset was from the CONVERGE consortium (CONVERGE consortium*, 2015) (referred as Chinese sample). Briefly, 5303 Chinese women with recurrent MDD and 5337 controls were recruited and low-coverage whole-genome sequencing was conducted. After stringent quality controls, 6,242,619 SNPs were remained and a linear mixed model was used to test the association between the SNPs and MDD. More detailed information about sample ascertainment, diagnosis, genotyping quality control, and statistical analyses can be found in the original studies (CONVERGE consortium*, 2015; Ripke et al., 2012).

For the meta-analysis, we first extracted the SNP associations from the Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium (Ripke et al., 2012) and the CONVERGE consortium (CONVERGE consortium*, 2015). Meta-analysis was then performed using MENTAL program (Willer et al., 2010), which uses an inverse-weighted fixed-effects model.

2.2. Brain eQTL data for integrative analysis

To elucidate if gene expression level in human brain is associated with genetic variation, Myers et al. systematically investigated the correlation between brain-expressed genes and whole-genome SNP genotypes (Myers et al., 2007). In brief, 193 neuropathologically normal (i.e., without known neurological diseases and had no clinical history of stroke, cerebrovascular disease and Lewy bodies) human brain samples (cortex) were used in the study of Myers et al. All of the subjects were of European ancestry. Myers et al. first measured whole-transcriptome expression data using the Illumina HumanRefseq-8 Expression BeadChip platforms. They then performed whole-genome genotyping by utilizing the Affymetrix GeneChip Human Mapping 500 K Array Set. Finally, they conducted allelic test of association to determine the potential association between each transcript and SNP through using linear regression. The association results were separated into *cis* and *trans* according to the genomic distance between the tested SNP and transcript. If the test SNP is located within the transcript, or located within 1 Mb of the 5′ or 3′ end of the transcript, this SNP was defined as *Cis* SNPs. SNPs that did not meet *cis* criteria were defined as *trans* SNPs. For more detailed information about sample ascertainment, gene expression measurement, SNP genotyping, and statistical analyses, please refer to the original study of Myers et al. (Myers et al., 2007).

2.3. Integrative analysis (Sherlock) of MDD GWAS data and brain eQTL

Numerous disease-associated genetic variants have been identified by GWAS (Hindorff et al., 2009a; Welter et al., 2014). However, how these identified genetic variants contribute to disease risk remains elusive. As most of the disease- and trait-associated genetic variants are located in noncoding regions (Hindorff et al., 2009b), elucidating the functional effects of these variants is a major challenge in human genetics. A growing body of evidence strongly suggests that dysregulation of gene expression play a crucial role in the pathogenesis of MDD (Bernard et al., 2010; Bunney et al., 2015). Considering that the genome-wide significant variants identified by the CONVERGE consortium are located in noncoding region, it is possible that these risk variants perturb the expression of MDD-associated genes rather than affect the protein

structure and function. To explore genes whose expression changes may contribute to MDD risk, we utilized a statistical framework (named *Sherlock*) developed by He et al. (He et al., 2013). *Sherlock* infers potential disease-associated genes through systematically integrating eQTL data and SNP associations from GWAS, with the underlying assumption that the expression alteration of a specific gene may influence disease risk. For a given gene, there may be several genetic variants (including *cis,* i.e., close to the gene, and *trans,* i.e., distant to the gene or on different chromosomes) in the genome that act together to regulate the expression level of this gene (these expression-associated SNPs are called eSNPs). Genotype difference at any of these eSNPs will influence the expression level of this gene, which could in turn affect the disease risk. Therefore, combined evidence from both eQTL and GWAS findings are indicative of association with disease. The *Sherlock* statistical inference includes three steps: First, through using the whole genome eQTL data from a disease-associated tissue (brain tissues were used in this study), *Sherlock* identifies all eSNPs of each gene. Second, for each eSNP, *Sherlock* assesses its association with disease using the GWAS data. Based on the association significance between the eSNP and the disease, *Sherlock* scores each eSNP. The scoring systems are as follows: (i) A positive score would be given if the eSNP of a specific gene is also associated with disease in GWAS; (ii) A negative score would be assigned if the eSNP of this gene is not associated with disease; (iii) Association only in GWAS (i.e., non-eSNPs) does not alter the score. The total score of a gene increases along with the increase of the number of SNPs with combined evidence (SNPs that are associated with certain disease and expression simultaneously). Finally, *Sherlock* infers gene-disease associations through using the Bayes statistical framework (He et al., 2013). Bayes factor (BF, the probability of the observed data under a specific model) is an important indicator that evaluates evidence supporting that the gene is associated versus not associated with the disease. For a given gene, the LBF (logarithm of BF) of each putative eSNP was calculated and the sum of LBFs of all SNPs constitutes the final LBF score for the gene. The value of the LBF score of a gene reflects the strength of evidence (i.e., a larger LBF represents higher probability that this gene is associated with the disease). *P* values of the LBF were calculated with simulation using the Bayes/non-Bayes compromise (Servin and Stephens, 2007) and false discovery rate (FDR) was calculated using the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995) at a *P* threshold of 10^{-5}. For more detailed information about the statistical model and the underlying algorithm, please refer to the original study of He et al. (He et al., 2013).

2.4. eQTL datasets for replication analysis

In addition to brain eQTL data from Myers et al. (which was used at the discovery phase in our study, i.e., *Sherlock* integrative analysis) (Myers et al., 2007), we also used other available well-characterized expression datasets to validate eQTL results. A brief description of the gene expression datasets from brain tissues and blood is provided below, and more detailed information can be found in the original studies (Ramasamy et al., 2014; The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans, 2015; Westra et al., 2014). (i) The Brain eQTL Almanac (Braineac) (Ramasamy et al., 2014), which contains genome-wide genotype data and whole transcriptome expression data (included 10 brain regions) of 134 normal human subjects (i.e., without neuropathological or neuropsychiatric diagnosis) (Ramasamy et al., 2014). After stringent quality control, eQTL analyses were performed to detect the association between genetic variants (SNPs) and gene expression. (ii) The Genotype-Tissue Expression project (GTEx) (The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans, 2015; The Genotype-Tissue Expression (GTEx) project, 2013). The goal of the GTEx consortium is to provide a resource with which to study human gene expression and regulation and its relationship to genetic variations. It contains genome-wide genetic information and gene expression from a diverse set of human tissues, including the different brain tissues from amygdala, anterior cingulate cortex, cortex, hippocampus, and so on. Tissues from different brain regions, including amygdala, anterior cingulate cortex, cortex, hippocampus (and etc) were also collected. As of January 2016, 7333 samples from 449 postmortem donors have been collected for the GTEx project. Expression data of these samples were analyzed using a RNA sequencing (RNA-seq)–based gene expression approach. (iii) Blood eQTL data (Westra et al., 2014). Westra et al. conducted a large-scale eQTL meta-analysis in non-transformed peripheral blood samples from 5311 individuals (Westra et al., 2014). Whole-genome eQTL data (peripheral blood) of 5311 samples from 7 cohorts were obtained and meta-analyzed. They identified numerous genetic variants that are robustly associated with gene expression in blood. More detailed information about sample description, genotyping, quality control, and statistical analyses can be found in the original paper (Westra et al., 2014).

2.5. Expression analysis of SLC25A37 in brain tissues of MDD patients

As the *Sherlock* integrative analysis suggested that the expression level of *SLC25A37* may contribute to MDD risk, we examined the expression of *SLC25A37* in brain tissues from MDD subjects and controls (Duric et al., 2010). To identify the dysregulated genes in hippocampus of MDD subjects, Duric et al. (Duric et al., 2010) collected postmortem brain tissues (hippocampus) from 21 MDD individuals and 18 controls, and performed a whole-genome expression profiling using the microarray method. MDD cases and healthy controls were matched for age, gender, tissue pH and postmortem interval. More detailed information about human subjects (including MDD diagnosis), tissue preparation, microarray analysis and statistical analyses can be found in the original study (Duric et al., 2010).

2.6. Expression analysis of SLC25A37 in blood of MDD patients (Chinese sample)

To have an independent validation of the result, we measured *SLC25A37* expression in peripheral blood from MDD patients and healthy controls. We recruited 50 unrelated MDD subjects (unmedicated) and 50 controls (age and sex-matched) from Shanghai Mental Health Center. All of the subjects are Han Chinese. MDD diagnosis was based on the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria. After providing written informed consent, each subject was diagnosed with standard diagnostic assessments by two experienced doctors. Assessments of the Hamilton Rating Scale for Depression -17 (HRSD-17) were conducted independently by two experienced psychiatrists (interrater reliability, kappa = 0.84). Demographic data on age, sex, smoking status, alcoholic abuse, duration of illness prior to admission, number of episode, family history of mood disorders was collected (Supplementary Table 1). This study was approved by Institutional Review Boards of Shanghai Mental Health Center and was performed in accordance with the guidelines laid out in the Declaration of Helsinki as revised in 1989.

On admission, 20 ml peripheral blood of fasting patients and healthy controls were collected between 07:00 a.m. and 09:00 a.m., to avoid potential diurnal influence. Total RNA was extracted from 10 ml peripheral blood samples using the QIAamp RNA blood Mini

Table 1
SNPs with *P*-values less than 5×10^{-7} in the combined sample (14,543 MDD cases and 14,856 controls).

SNP id (Chr)	Chr:Pos[a]	Nearest gene	Location	A_{12}[b]	OR[c] (CEU)	P[c] (CEU)	OR[d] (CHB)	P[d] (CHB)	OR[e] (Combined)	P[e] (Combined)
rs12549100	8:23461175	*SLC25A37*	Intergenic	A/G	0.939	4.99×10^{-3}	0.859	2.43×10^{-6}	0.910	1.05×10^{-7}
rs6983724	8:23457587	*SLC25A37*	Intergenic	A/G	1.06	6.36×10^{-3}	1.16	2.95×10^{-6}	1.10	1.65×10^{-7}
rs11995691	8:23469198	*SLC25A37*	Intergenic	A/G	1.06	6.84×10^{-3}	1.16	3.16×10^{-6}	1.10	1.91×10^{-7}
rs4862792	4:188201350	*LOC339975*	Intergenic	T/G	1.15	2.09×10^{-5}	1.13	1.94×10^{-3}	1.10	1.99×10^{-7}
rs11995896	8:23472742	*SLC25A37*	Intergenic	A/C	0.941	6.66×10^{-3}	0.860	3.50×10^{-6}	0.912	2.01×10^{-7}
rs7647854	3:184876783	*EHHADH-AS1*	Intergenic	A/G	0.860	6.51×10^{-7}	0.93	2.15×10^{-2}	0.892	2.69×10^{-7}
rs16836496	2:155176299	*GALNT13*	Intron	T/C	1.10	8.19×10^{-4}	1.13	1.35×10^{-4}	1.11	4.08×10^{-7}

[a] Chromosome position, based on hg19.
[b] A_{12}, Reference and alternative alleles (Odds ratios are based on reference allele).
[c] Odd ratios and *P* values are from European sample (The Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium) (Ripke et al., 2012).
[d] Odd ratios and *P* values are from Chinese sample (The CONVERGE consortium) (CONVERGE consortium*, 2015).
[e] Odd ratios and *P* values of the combined samples.

Kit (Qiagen, Chatsworth, California, USA). After treating with DNase (Qiagen, Chatsworth, California, USA) (to remove the potential DNA contamination), complementary DNA (cDNA) was synthesized by incubating DNase-treated total RNA with omniscript reverse transcription reagents (Qiagen, Chatsworth, California, USA) and a random primer according to the manufacturer's instruction. Real-time quantitative PCR were performed as previously described (Luo et al., 2012, 2013) using the SYBR Green master mix and the Bio-Rad CFX Real-Time PCR Systems (BioRad). The specificity of product was assessed by agrose gel electrophoresis, followed with melting curve analysis. The expression of *SLC25A37* in each sample was normalized to the expression of *GAPDH*. The analysis of quantitative PCR data was based on the ΔC_t values and fold change was determined by $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). All reactions were run in triplicate and statistical significance was evaluated by *student's* t-test.

2.7. Expression analysis of SLC25A37 in blood of MDD patients (European sample)

In addition to Chinese sample, we also explored the expression of *SLC25A37* in peripheral blood of MDD patients and healthy controls using an independent European sample (Jansen et al., 2016). To investigate if dysregulation of gene expression play a role in MDD, Jansen et al. performed a large-scale expression study (Jansen et al., 2016). Briefly, they measured gene expression in peripheral blood from 1848 subjects from The Netherlands Study of Depression and Anxiety. These subjects were divided into three groups, including current MDD (N = 882), remitted MDD (N = 635) and control (N = 331) groups. More detailed information about human subjects, tissue preparation, gene expression measurement, quality control and statistical analyses can be found in the original study (Jansen et al., 2016).

3. Results

3.1. Meta-analysis of MDD GWAS

Meta-analysis of GWAS datasets from the Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium (Ripke et al., 2012) and the CONVERGE consortium (CONVERGE consortium*, 2015) revealed that no SNPs reached genome-wide significant level ($P < 5.0 \times 10^{-8}$). However, we detected several interesting SNPs that have the same direction of effect in both European and Chinese samples. Table 1 lists the SNPs that have *P*-values less than 5×10^{-7} in the combined sample (14,543 MDD cases and 14,856 controls). These SNPs may represent promising MDD risk variants. Further work will be required to verify if these SNPs are associated with MDD.

3.2. Integration of MDD GWAS and brain eQTL identified SLC25A37 as a MDD risk gene

Meta-analysis results from the combined sample (European and Chinese) was used as input for further *Sherlock* integrative analysis. Through systematic integration of brain eQTL (Myers et al., 2007) and SNP associations from meta-analysis of the two large GWAS of MDD (CONVERGE consortium*, 2015; Ripke et al., 2012), *SLC25A37* (8p21.1) showed the most significant association with MDD (LBF = 6.79, $P_{sher} = 2.2 \times 10^{-6}$) (Table 2). One eSNP of *SLC25A37* (rs6983724) showed significant association with *SLC25A37* expression ($P_{eQTL} = 1.19 \times 10^{-9}$) and suggestive evidence for association with *MDD* ($P = 1.65 \times 10^{-7}$) (Table 2). We noticed that individuals with AA genotype at rs6983724 have lower *SLC25A37* expression compared with GG carriers (Fig. 1). Intriguingly, A allele of SNP rs6983724 is the risk allele in both European (OR = 1.06) and Han Chinese (OR = 1.16) populations (Table 1). The top five genes

Table 2
Integrative analysis (*Sherlock*) of MDD GWAS and brain eQTL reveals that *SLC25A37* had the most significant association with MDD.

Gene symbol	LBF[a]	*P*-value[b]	Supporting SNP[c] (*cis* or *trans*)	P_{eQTL}[d]	P_{GWAS}[e]
SLC25A37	6.79	2.22×10^{-6}	rs6983724 (*cis*)	1.19×10^{-9}	1.65×10^{-7}
PGAM2	5.44	1.33×10^{-5}	rs17138095 (*trans*)	5.33×10^{-6}	1.33×10^{-4}
MDM1	4.98	2.22×10^{-5}	rs4478240 (*trans*)	9.00×10^{-7}	5.81×10^{-5}
PTPRR	4.82	3.33×10^{-5}	rs13132638 (*trans*)	1.05×10^{-7}	8.24×10^{-4}
SNCA	4.55	5.99×10^{-5}	rs929313 (*trans*)	1.11×10^{-6}	1.92×10^{-5}

[a] LBF (Log Bayes Factors for each gene) assesses whether the combined evidence from MDD GWAS and brain expression studies support a gene being associated with MDD. High Bayes factors suggest higher probability that the gene is associated with MDD.
[b] *P*-value from *Sherlock* integrative analysis.
[c] SNP with the highest LBF score.
[d] *P*-value for eQTL SNP from the gene expression study (Myers et al., 2007).
[e] *P*-value for eQTL SNP from the meta-analysis of MDD GWAS (CONVERGE consortium*, 2015; Ripke et al., 2012). More detailed descriptions about the LBF can be found in our previous study (Luo et al., 2015). Cis SNPs were defined as follows: SNPs located within gene, SNPs located upstream of the gene (within 1 Mb) (5′), or SNPs located within 1 Mb downstream of the gene (3′).

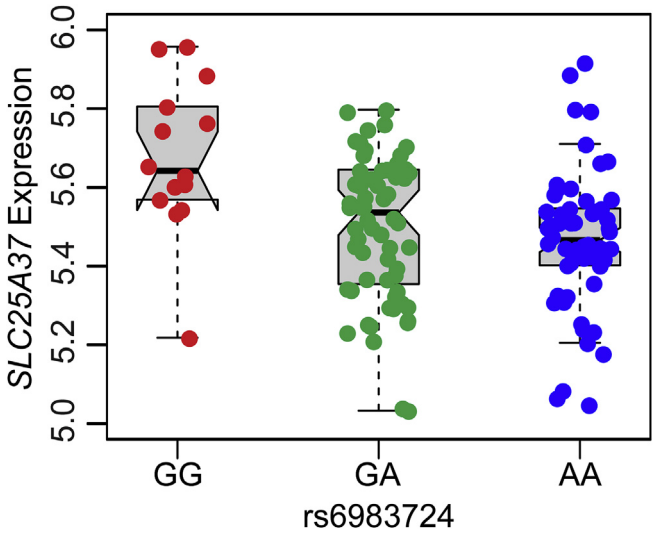

Fig. 1. SNP rs6983724 is significantly associated with *SLC25A37* expression in human brain ($P = 2.6 \times 10^{-3}$). Expression data was extracted from the Braineac (Ramasamy et al., 2014), which contained brain tissues of 134 normal subjects. Of note, the individuals with AA genotype have lower *SLC25A37* expression level.

from *Sherlock* integrative analysis are listed in Table 2.

3.3. SNP rs6983724 is associated with SLC25A37 expression in independent datasets

Sherlock integrative analysis suggested that *SLC25A37* may be a novel MDD risk gene. Interestingly, we found that SNP rs6983724 is significantly associated with *SLC25A37* expression in human brain, suggesting that rs6983724 may confer risk of MDD through affecting *SLC25A37* expression. To further validate the association between rs6983724 and *SLC25A37* expression, we examined three independent expression datasets. Two of the expression datasets (the Braineac (Ramasamy et al., 2014) and the GTEx (The Genotype-Tissue Expression (GTEx) project, 2013)) used human brain tissues, and one expression dataset used blood sample (Westra et al., 2014). We found that rs6983724 is significantly associated with *SLC25A37* expression ($P = 2.6 \times 10^{-3}$) in Braineac dataset (N = 134 subjects) (Fig. 1). In GTEx dataset, rs6983724 is also significantly associated with *SLC25A37* expression ($P < 0.05$) in cortex and anterior cingulate cortex (Supplementary Fig. 1). Consistently, rs6983724 is also highly associated with *SLC25A37* expression ($P = 1.06 \times 10^{-167}$) in the peripheral blood sample dataset (N = 5311 individuals) from Westra et al. (Westra et al., 2014). SNP rs6983724 is located in ~28 kb downstream of *SLC25A37* (Supplementary Fig. 2), according to the criteria described in previous papers (Bryois et al., 2014; He et al., 2013; Li and Deng, 2010; Wittkopp, 2005), rs6983724 is a cis eQTL of *SLC25A37*. These consistent results suggested that rs6983724 is an authentic eQTL for *SLC25A37* gene.

3.4. Significant downregulation of SLC25A37 in the hippocampus and peripheral blood of MDD patients

Our *Sherlock* integrative analysis suggested that the expression alterations of *SLC25A37* may contribute to MDD risk. To further verify this, we tested whether *SLC25A37* expression in hippocampus was dysregulated in MDD patients using expression data from Duric et al. (Duric et al., 2010). We found that *SLC25A37* expression is significantly down-regulated in MDD patients compared with normal controls (decreased by 38%) ($P = 3.49 \times 10^{-3}$) (Supplementary Fig. 3). We further examined if *SLC25A37* is differentially expressed in peripheral blood of MDD cases and controls using the data from a recent large-scale MDD expression study (Jansen et al., 2016). Compared with controls, *SLC25A37* expression is significantly down-regulated in current MDD cases (N = 882 cases and 331 controls, $P = 3.06 \times 10^{-3}$) and remitted MDD cases (N = 635 cases and 331 controls, $P = 1.13 \times 10^{-3}$). When current and remitted MDD cases were combined (N = 1157 cases and 331 controls), a more significant *P* value was observed ($P = 7.90 \times 10^{-4}$). Finally, we validated the significant down-regulation of *SLC25A37* in Chinese sample using quantitative PCR (down-regulated by 55%) ($P = 2.66 \times 10^{-13}$) (Fig. 2). These consistent results indicated that *SLC25A37* expression is significantly down-regulated in MDD cases. Taken together, these convergent lines of evidence suggested that *SLC25A37* may represent a novel MDD susceptibility gene.

4. Discussion

MDD is one of the most prevalent mental disorders. Though the heritability of MDD is relatively high (around 0.32) (Lubke et al., 2012; Sullivan et al., 2000), only very limited MDD risk variants (or genes) have been identified by recent large-scale GWAS. The conundrum of genetic studies suggest the complex genetic architecture of MDD. As stated in the work of Ripke et al. (Ripke et al., 2012), sample size may be an important factor and new risk loci may be identified with the increase of sample size. Consistent with this, Hyde et al. recently reported the largest GWAS of MDD (Hyde et al., 2016). A total of 75,607 MDD cases and 231,747 controls were included in the first stage (discovery), and 45,773 MDD cases and 106,354 controls were included in the second stage (replication). Hyde et al. successfully identified 15 novel genetic loci for MDD through combing the results from discovery and replication phases. The genes located near the identified risk variants including *TEME161B*, *MEF2C*, *VRK2*, *L3MBTL2*, *NEGR1*, *RERE*, *HACE1*, *LIN28B*, *SORCS3*, *OLFM4*, *PAX5*, *MEIS2*, *TMCO5A*, *RSRC1*, *MLF1*, *SLC6A15*, *KIAA0020 and RFX3*. These important findings further indicate the pivotal role of sample size in genetic study of MDD.

In this study, we performed a meta-analysis through combing two large GWAS datasets of MDD. Though no genome-wide

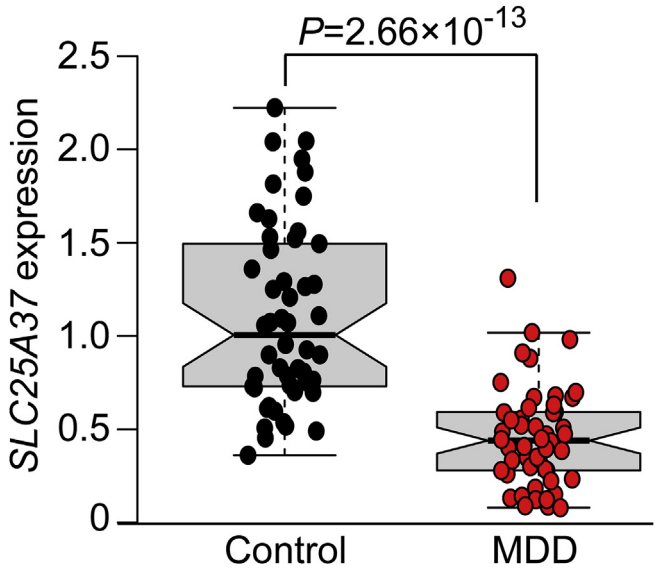

Fig. 2. *SLC25A37* is significantly down-regulated in peripheral blood of MDD patients. Compared with healthy controls, *SLC25A37* expression was significantly down-regulated in MDD cases.

significant variants were observed, we identified several promising candidate SNPs that have the same direction of effect in both European and Chinese samples. These SNPs may represent promising candidate variants for MDD and genome-wide significant associations may be observed with the increase of sample size. Four of the seven top SNPs are located 8p21.1 (Supplementary Fig. 2), suggesting this region may harbor novel MDD risk variants. In addition to the identification of a potential risk variant (rs6983724) for MDD, our study also suggested that integrative analysis may provide a useful way to mine the missing heritability of complex diseases. Most of the disease-associated genetic variants identified by GWAS are located in non-coding region, suggesting that these risk variants manifest their effects through regulating gene expression. Thus, characterization of the regulatory architecture of these variants is important for interpreting GWAS loci and understanding basic biology. Expression quantitative trait locus (eQTL) is one of the most common methods used to dissect the effects of non-coding genetic variation. Though a large amount of disease and trait-associated genetic variants have been identified (Welter et al., 2014) and numerous tissue-specific eQTL databases (The Genotype-Tissue Expression (GTEx) project, 2013; Westra et al., 2014) are available, how to systematically integrate these two different data types in genetic study remains a major challenge. Recently, He et al. developed a novel method (named *Sherlock*) to detect gene-disease associations by matching patterns of expression QTL and GWAS (He et al., 2013). Through using this integrative method (*Sherlock*), we identified *ZNF323* as a novel schizophrenia risk gene (Luo et al., 2015). More important, we successfully replicated the association between *ZNF323* and schizophrenia in independent samples, suggesting that *Sherlock* integrative analysis is powerful to detect disease-associated genes through integrating eQTL and GWAS.

Through using *Sherlock* integrative analysis, we showed that *SLC25A37* may be a novel risk gene for MDD. First, among the 7 top SNPs ($P < 5.0 \times 10^{-7}$) identified by meta-analysis (Table 1), four of them (rs6983724, rsrs12549100, rs11995691, and rs11995896) (Table 1) are located in the downstream of *SLC25A37*. Linkage disequilibrium analysis using genotype data (Europeans) from the 1000 Genomes project (Abecasis et al., 2010) indicated that these four SNPs are highly linked ($r^2 > 0.95$) (Supplementary Fig. 4). In fact, *SLC25A37* is the nearest gene for the these four SNPs. Second, systematic integration brain eQTL and MDD GWAS revealed that *SLC25A37* had the most significant association with MDD, further supporting that *SLC25A37* may be a MDD risk gene. SNP rs6983724 is significantly associated with *SLC25A37* expression and suggestive association with MDD. Intriguingly, among the four SNPs (rs6983724, rsrs12549100, rs11995691, and rs11995896) that located in the downstream of *SLC25A37*, rs6983724 has the nearest distance to *SLC25A37* (Supplementary Fig. 2). Compared with GG genotype carriers, individuals with AA genotype at SNP rs6983724 showed a reduced expression of *SLC25A37* (Fig. 1). We validated the association between rs6983724 and *SLC25A37* expression in independent datasets. Third, expression analyses indicated that *SLC25A37* is significantly down-regulated in MDD patients. We also verified the down-regulation of *SLC25A37* expression in independent populations. Finally, SNP rs6983724 showed significant association with MDD in both European and Chinese samples. Of note, we found that the A allele of rs6983724 is the risk allele in both Chinese (OR = 1.16) and European (OR = 1.06) samples. The strengths of association significance between rs6983724 and MDD increases when Chinese and European samples were combined, strongly suggesting that rs6983724 may represent a promising risk variant for MDD. These convergent evidence suggest that *SLC25A37* may be a new MDD risk gene.

SLC25A37 (Solute Carrier Family 25 Member 37) encodes mitochondrial iron transporter (Mitoferrin-1) that is essential for erythroid iron assimilation (Shaw et al., 2006). It functions as an essential iron importer that imports ferrous iron from the intermembrane space of the mitochondria to the mitochondrial matrix for the synthesis of heme groups and Fe–S clusters (Chen et al., 2009; Paradkar et al., 2009). Expression analysis revealed that *SLC25A37* is widely expressed in central nervous system (Haitina et al., 2006), suggesting that it may play a role in brain development and function. Recent studies revealed that *SLC25A37* is important for many biological processes, including erythroid heme biosynthesis (Chen et al., 2010; Shaw et al., 2006), fertility (Metzendorf and Lind, 2010), body size and movement (Ren et al., 2012). SLC25A37 is physically interacted with ABCB10, a transporter that also plays an important role in heme biosynthesis (Chen et al., 2009). Through forming complex, ABCB10 can enhance SLC25A37 stability and function, and loss of *Abcb10* greatly impacted the heme biosynthesis (Yamamoto et al., 2014). Visconte et al. showed that *SLC25A37* is regulated by SF3B1 (Visconte et al., 2015), a member of the RNA slicing machinery. Mutation of *SF3B1* resulted in overexpression of *SLC25A37*, which eventually lead to iron overload. In addition, abnormal expression of *SLC25A37* was observed in human diseases such as erythropoietic protoporphyria (Wang et al., 2011) and reduction of Mitoferrin resulted in abnormal development (Ren et al., 2012). Interestingly, a recent study reported by Korkmaz et al. showed that among MDD patients, 22% was diagnosed with anemia (Korkmaz et al., 2015), strongly suggesting that *SLC25A37* may have pivotal role in MDD. We noticed that Mitoferrin-1 is an important component of mitochondrion, the central machinery to produce the energy currency (ATP) of the cell. Accumulating evidence suggested that mitochondrial dysfunction may have pivotal role in MDD (Cai et al., 2015; Chang et al., 2015; Gardner and Boles, 2010; Klinedinst and Regenold, 2015; Rezin et al., 2009; Tobe, 2013). Consistently, the CONVERGE consortium also found that one of the two genome-wide loci is located to *SIRT1*, a gene involved in mitochondrial biogenesis (Gerhart-Hines et al., 2007). These evidence imply that mitochondrial dysfunction may be associated with abnormal brain function and MDD. Further work is needed to validate our findings in a larger sample or in an independent set of samples.

Contributors

XJL designed and directed the research. XJL, YXH, LH, DFZ, and CZ performed the research. XJL, YXH, and CZ analyzed the data. XJL, CZ, YRF, and YGY wrote the paper. All authors read and approved the final manuscript.

Funding

X.J.L was supported by the 100 Talents Program (*BaiRenJiHua*) of the Kunming Institute of Zoology, Chinese Academy of Sciences, and Project of Thousand Youth Talents of China. Y.G.Y was supported by the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (XDB02020003). C.Z. was supported by the National Natural Science Foundation of China (81471358) and the Shanghai Municipal Education Commission—Gaofeng Clinical Medicine Grant Support (20152530).

Competing financial interests

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

Acknowledgements

We thank Di Huang for her technical support. We also thank the research participants for their contributions.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jpsychires.2016.09.011.

References

Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., McVean, G.A., 2010. A map of human genome variation from population-scale sequencing. Nature 467, 1061–1073.

Angst, F., Stassen, H.H., Clayton, P.J., Angst, J., 2002. Mortality of patients with mood disorders: follow-up over 34-38 years. J. Affect. Disord. 68, 167–181.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B 57, 289–300.

Bernard, R., Kerman, I.A., Thompson, R.C., Jones, E.G., Bunney, W.E., Barchas, J.D., Schatzberg, A.F., Myers, R.M., Akil, H., Watson, S.J., 2010. Altered expression of glutamate signaling, growth factor, and glia genes in the locus coeruleus of patients with major depression. Mol. Psychiatry 16, 634–646.

Bryois, J., Buil, A., Evans, D.M., Kemp, J.P., Montgomery, S.B., Conrad, D.F., Ho, K.M., Ring, S., Hurles, M., Deloukas, P., Davey Smith, G., Dermitzakis, E.T., 2014. Cis and trans effects of human genomic variants on gene expression. PLoS Genet. 10, e1004461.

Bunney, B.G., Li, J.Z., Walsh, D.M., Stein, R., Vawter, M.P., Cartagena, P., Barchas, J.D., Schatzberg, A.F., Myers, R.M., Watson, S.J., Akil, H., Bunney, W.E., 2015. Circadian dysregulation of clock genes: clues to rapid treatments in major depressive disorder. Mol. Psychiatry 20, 48–55.

Cai, N., Chang, S., Li, Y., Li, Q., Hu, J., Liang, J., Song, L., Kretzschmar, W., Gan, X., Nicod, J., Rivera, M., Deng, H., Du, B., Li, K., Sang, W., Gao, J., Gao, S., Ha, B., Ho, H.Y., Hu, C., Hu, Z., Huang, G., Jiang, G., Jiang, T., Jin, W., Li, G., Lin, Y.T., Liu, L., Liu, T., Liu, Y., Lu, Y., Lv, L., Meng, H., Qian, P., Sang, H., Shen, J., Shi, J., Sun, J., Tao, M., Wang, G., Wang, J., Wang, L., Wang, X., Yang, H., Yang, L., Yin, Y., Zhang, J., Zhang, K., Sun, N., Zhang, W., Zhang, X., Zhang, Z., Zhong, H., Breen, G., Marchini, J., Chen, Y., Xu, Q., Xu, X., Mott, R., Huang, G.J., Kendler, K., Flint, J., 2015. Molecular signatures of major depression. Curr. Biol. 25, 1146–1156.

Chang, C.C., Jou, S.H., Lin, T.T., Lai, T.J., Liu, C.S., 2015. Mitochondria DNA change and oxidative damage in clinically stable patients with major depressive disorder. PLoS One 10, e0125855.

Chen, W., Dailey, H.A., Paw, B.H., 2010. Ferrochelatase forms an oligomeric complex with mitoferrin-1 and Abcb10 for erythroid heme biosynthesis. Blood 116, 628–630.

Chen, W., Paradkar, P.N., Li, L., Pierce, E.L., Langer, N.B., Takahashi-Makise, N., Hyde, B.B., Shirihai, O.S., Ward, D.M., Kaplan, J., Paw, B.H., 2009. Abcb10 physically interacts with mitoferrin-1 (Slc25a37) to enhance its stability and function in the erythroid mitochondria. Proc. Natl. Acad. Sci. U. S. A. 106, 16263–16268.

CONVERGE consortium*, 2015. Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature 523, 588–591.

Duric, V., Banasr, M., Licznerski, P., Schmidt, H.D., Stockmeier, C.A., Simen, A.A., Newton, S.S., Duman, R.S., 2010. A negative regulator of MAP kinase causes depressive behavior. Nat. Med. 16, 1328–1332.

Ferrari, A.J., Charlson, F.J., Norman, R.E., Patten, S.B., Freedman, G., Murray, C.J., Vos, T., Whiteford, H.A., 2013. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS Med. 10, e1001547.

Flint, J., Kendler, K.S., 2014. The genetics of major depression. Neuron 81, 484–503.

Gardner, A., Boles, R.G., 2010. Beyond the serotonin hypothesis: mitochondria, inflammation and neurodegeneration in major depression and affective spectrum disorders. Prog. Neuropsychopharmacol. Biol. Psychiatry 35, 730–743.

Gerhart-Hines, Z., Rodgers, J.T., Bare, O., Lerin, C., Kim, S.H., Mostoslavsky, R., Alt, F.W., Wu, Z., Puigserver, P., 2007. Metabolic control of muscle mitochondrial function and fatty acid oxidation through SIRT1/PGC-1alpha. EMBO J. 26, 1913–1923.

Greenberg, P.E., Fournier, A.A., Sisitsky, T., Pike, C.T., Kessler, R.C., 2015. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). J. Clin. Psychiatry 76, 155–162.

Haitina, T., Lindblom, J., Renstrom, T., Fredriksson, R., 2006. Fourteen novel human members of mitochondrial solute carrier family 25 (SLC25) widely expressed in the central nervous system. Genomics 88, 779–790.

Hasin, D.S., Goodwin, R.D., Stinson, F.S., Grant, B.F., 2005. Epidemiology of major depressive disorder: results from the national epidemiologic survey on alcoholism and related conditions. Arch. Gen. Psychiatry 62, 1097–1106.

He, X., Fuller, C.K., Song, Y., Meng, Q., Zhang, B., Yang, X., Li, H., 2013. Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. Am. J. Hum. Genet. 92, 667–680.

Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., Manolio, T.A., 2009a. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. U. S. A. 106, 9362–9367.

Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., Manolio, T.A., 2009b. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. U. S. A. 106, 9362–9367.

Hyde, C.L., Nagle, M.W., Tian, C., Chen, X., Paciga, S.A., Wendland, J.R., Tung, J.Y., Hinds, D.A., Perlis, R.H., Winslow, A.R., 2016. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. Nat. Genet. 48, 1031–1036.

Jansen, R., Penninx, B.W., Madar, V., Xia, K., Milaneschi, Y., Hottenga, J.J., Hammerschlag, A.R., Beekman, A., van der Wee, N., Smit, J.H., Brooks, A.I., Tischfield, J., Posthuma, D., Schoevers, R., van Grootheest, G., Willemsen, G., de Geus, E.J., Boomsma, D.I., Wright, F.A., Zou, F., Sun, W., Sullivan, P.F., 2016. Gene expression in major depressive disorder. Mol. Psychiatry 21, 339–347. http://dx.doi.org/10.1038/mp.2015.57. Advance online publication, 26 May 2015.

Kessler, R.C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K.R., Rush, A.J., Walters, E.E., Wang, P.S., 2003. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 289, 3095–3105.

Klinedinst, N.J., Regenold, W.T., 2015. A mitochondrial bioenergetic basis of depression. J. Bioenerg. Biomembr. 47, 155–171.

Kohli, M.A., Lucae, S., Saemann, P.G., Schmidt, M.V., Demirkan, A., Hek, K., Czamara, D., Alexander, M., Salyakina, D., Ripke, S., Hoehn, D., Specht, M., Menke, A., Hennings, J., Heck, A., Wolf, C., Ising, M., Schreiber, S., Czisch, M., Muller, M.B., Uhr, M., Bettecken, T., Becker, A., Schramm, J., Rietschel, M., Maier, W., Bradley, B., Ressler, K.J., Nothen, M.M., Cichon, S., Craig, I.W., Breen, G., Lewis, C.M., Hofman, A., Tiemeier, H., van Duijn, C.M., Holsboer, F., Muller-Myhsok, B., Binder, E.B., 2011. The neuronal transporter gene SLC6A15 confers risk to major depression. Neuron 70, 252–265.

Korkmaz, S., Yildiz, S., Korucu, T., Gundogan, B., Sunbul, Z.E., Korkmaz, H., Atmaca, M., 2015. Frequency of anemia in chronic psychiatry patients. Neuropsychiatr. Dis. Treat. 11, 2737–2741.

Lesch, K.P., 2004. Gene-environment interaction and the genetics of depression. J. Psychiatry. Neurosci. 29, 174–184.

Lewis, C.M., Ng, M.Y., Butler, A.W., Cohen-Woods, S., Uher, R., Pirlo, K., Weale, M.E., Schosser, A., Paredes, U.M., Rivera, M., Craddock, N., Owen, M.J., Jones, L., Jones, I., Korszun, A., Aitchison, K.J., Shi, J., Quinn, J.P., Mackenzie, A., Vollenweider, P., Waeber, G., Heath, S., Lathrop, M., Muglia, P., Barnes, M.R., Whittaker, J.C., Tozzi, F., Holsboer, F., Preisig, M., Farmer, A.E., Breen, G., Craig, I.W., McGuffin, P., 2010. Genome-wide association study of major recurrent depression in the U.K. population. Am. J. Psychiatry 167, 949–957.

Li, H., Deng, H., 2010. Systems genetics, bioinformatics and eQTL mapping. Genetica 138, 915–924.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402–408.

Lopez, A.D., Mathers, C.D., Ezzati, M., Jamison, D.T., Murray, C.J., 2006. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet 367, 1747–1757.

Lubke, G.H., Hottenga, J.J., Walters, R., Laurin, C., de Geus, E.J., Willemsen, G., Smit, J.H., Middeldorp, C.M., Penninx, B.W., Vink, J.M., Boomsma, D.I., 2012. Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. Biol. Psychiatry 72, 707–709.

Luo, X.J., Deng, M., Xie, X., Huang, L., Wang, H., Jiang, L., Liang, G., Hu, F., Tieu, R., Chen, R., Gan, L., 2013. GATA3 controls the specification of prosensory domain and neuronal survival in the mouse cochlea. Hum. Mol. Genet. 22, 3609–3623.

Luo, X.J., Li, M., Huang, L., Nho, K., Deng, M., Chen, Q., Weinberger, D.R., Vasquez, A.A., Rijpkema, M., Mattay, V.S., Saykin, A.J., Shen, L., Fernandez, G., Franke, B., Chen, J.C., Chen, X.N., Wang, J.K., Xiao, X., Qi, X.B., Xiang, K., Peng, Y.M., Cao, X.Y., Li, Y., Shi, X.D., Gan, L., Su, B., 2012. The interleukin 3 gene (IL3) contributes to human brain volume variation by regulating proliferation and survival of neural progenitors. PLoS One 7, e50375.

Luo, X.J., Mattheisen, M., Li, M., Huang, L., Rietschel, M., Borglum, A.D., Als, T.D., van den Oord, E.J., Aberg, K.A., Mors, O., Mortensen, P.B., Luo, Z., Degenhardt, F., Cichon, S., Schulze, T.G., Nothen, M.M., Su, B., Zhao, Z., Gan, L., Yao, Y.G., 2015. 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. 41, 1294–1308.

Metzendorf, C., Lind, M.I., 2010. Drosophila mitoferrin is essential for male fertility: evidence for a role of mitochondrial iron metabolism during spermatogenesis. BMC. Dev. Biol. 10, 68.

Muglia, P., Tozzi, F., Galwey, N.W., Francks, C., Upmanyu, R., Kong, X.Q., Antoniades, A., Domenici, E., Perry, J., Rothen, S., Vandeleur, C.L., Mooser, V., Waeber, G., Vollenweider, P., Preisig, M., Lucae, S., Muller-Myhsok, B., Holsboer, F., Middleton, L.T., Roses, A.D., 2010. Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. Mol. Psychiatry 15, 589–601.

Myers, A.J., Gibbs, J.R., Webster, J.A., Rohrer, K., Zhao, A., Marlowe, L., Kaleem, M., Leung, D., Bryden, L., Nath, P., Zismann, V.L., Joshipura, K., Huentelman, M.J., Hu-Lince, D., Coon, K.D., Craig, D.W., Pearson, J.V., Holmans, P., Heward, C.B., Reiman, E.M., Stephan, D., Hardy, J., 2007. A survey of genetic human cortical gene expression. Nat. Genet. 39, 1494–1499.

Olchanski, N., McInnis Myers, M., Halseth, M., Cyr, P.L., Bockstedt, L., Goss, T.F.,

Howland, R.H., 2013. The economic burden of treatment-resistant depression. Clin. Ther. 35, 512–522.
Paradkar, P.N., Zumbrennen, K.B., Paw, B.H., Ward, D.M., Kaplan, J., 2009. Regulation of mitochondrial iron import through differential turnover of mitoferrin 1 and mitoferrin 2. Mol. Cell Biol. 29, 1007–1016.
Ramasamy, A., Trabzuni, D., Guelfi, S., Varghese, V., Smith, C., Walker, R., De, T., Coin, L., de Silva, R., Cookson, M.R., Singleton, A.B., Hardy, J., Ryten, M., Weale, M.E., 2014. Genetic variability in the regulation of gene expression in ten regions of the human brain. Nat. Neurosci. 17, 1418–1428.
Ren, Y., Yang, S., Tan, G., Ye, W., Liu, D., Qian, X., Ding, Z., Zhong, Y., Zhang, J., Jiang, D., Zhao, Y., Lu, J., 2012. Reduction of mitoferrin results in abnormal development and extended lifespan in Caenorhabditis elegans. PLoS One 7, e29666.
Rezin, G.T., Amboni, G., Zugno, A.I., Quevedo, J., Streck, E.L., 2009. Mitochondrial dysfunction and psychiatric disorders. Neurochem. Res. 34, 1021–1029.
Rietschel, M., Mattheisen, M., Frank, J., Treutlein, J., Degenhardt, F., Breuer, R., Steffens, M., Mier, D., Esslinger, C., Walter, H., Kirsch, P., Erk, S., Schnell, K., Herms, S., Wichmann, H.E., Schreiber, S., Jockel, K.H., Strohmaier, J., Roeske, D., Haenisch, B., Gross, M., Hoefels, S., Lucae, S., Binder, E.B., Wienker, T.F., Schulze, T.G., Schmal, C., Zimmer, A., Juraeva, D., Brors, B., Bettecken, T., Meyer-Lindenberg, A., Muller-Myhsok, B., Maier, W., Nothen, M.M., Cichon, S., 2010. Genome-wide association-, replication-, and neuroimaging study implicates HOMER1 in the etiology of major depression. Biol. Psychiatry 68, 578–585.
Ripke, S., Wray, N.R., Lewis, C.M., Hamilton, S.P., Weissman, M.M., Breen, G., Byrne, E.M., Blackwood, D.H., Boomsma, D.I., Cichon, S., Heath, A.C., Holsboer, F., Lucae, S., Madden, P.A., Martin, N.G., McGuffin, P., Muglia, P., Noethen, M.M., Penninx, B.P., Pergadia, M.L., Potash, J.B., Rietschel, M., Lin, D., Muller-Myhsok, B., Shi, J., Steinberg, S., Grabe, H.J., Lichtenstein, P., Magnusson, P., Perlis, R.H., Preisig, M., Smoller, J.W., Stefansson, K., Uher, R., Kutalik, Z., Tansey, K.E., Teumer, A., Viktorin, A., Barnes, M.R., Bettecken, T., Binder, E.B., Breuer, R., Castro, V.M., Churchill, S.E., Coryell, W.H., Craddock, N., Craig, I.W., Czamara, D., De Geus, E.J., Degenhardt, F., Farmer, A.E., Fava, M., Frank, J., Gainer, V.S., Gallagher, P.J., Gordon, S.D., Goryachev, S., Gross, M., Guipponi, M., Henders, A.K., Herms, S., Hickie, I.B., Hoefels, S., Hoogendijk, W., Hottenga, J.J., Iosifescu, D.V., Ising, M., Jones, I., Jones, L., Jung-Ying, T., Knowles, J.A., Kohane, I.S., Kohli, M.A., Korszun, A., Landen, M., Lawson, W.B., Lewis, G., Macintyre, D., Maier, W., Mattheisen, M., McGrath, P.J., McIntosh, A., McLean, A., Middeldorp, C.M., Middleton, L., Montgomery, G.M., Murphy, S.N., Nauck, M., Nolen, W.A., Nyholt, D.R., O'Donovan, M., Oskarsson, H., Pedersen, N., Scheftner, W.A., Schulz, A., Schulze, T.G., Shyn, S.I., Sigurdsson, E., Slager, S.L., Smit, J.H., et al., 2012. A mega-analysis of genome-wide association studies for major depressive disorder. Mol. Psychiatry 18, 497–511.
Servin, B., Stephens, M., 2007. Imputation-based analysis of association studies: candidate regions and quantitative traits. PLoS Genet. 3, e114.
Shaw, G.C., Cope, J.J., Li, L., Corson, K., Hersey, C., Ackermann, G.E., Gwynn, B., Lambert, A.J., Wingert, R.A., Traver, D., Trede, N.S., Barut, B.A., Zhou, Y., Minet, E., Donovan, A., Brownlie, A., Balzan, R., Weiss, M.J., Peters, L.L., Kaplan, J., Zon, L.I., Paw, B.H., 2006. Mitoferrin is essential for erythroid iron assimilation. Nature 440, 96–100.
Shi, J., Potash, J.B., Knowles, J.A., Weissman, M.M., Coryell, W., Scheftner, W.A., Lawson, W.B., DePaulo Jr., J.R., Gejman, P.V., Sanders, A.R., Johnson, J.K., Adams, P., Chaudhury, S., Jancic, D., Evgrafov, O., Zvinyatskovskiy, A., Ertman, N., Gladis, M., Neimanas, K., Goodell, M., Hale, N., Ney, N., Verma, R., Mirel, D., Holmans, P., Levinson, D.F., 2011. Genome-wide association study of recurrent early-onset major depressive disorder. Mol. Psychiatry 16, 193–201.
Subbarao, A., Rhee, S.H., Young, S.E., Ehringer, M.A., Corley, R.P., Hewitt, J.K., 2008. Common genetic and environmental influences on major depressive disorder and conduct disorder. J. Abnorm. Child. Psychol. 36, 433–444.
Sullivan, P.F., Daly, M.J., O'Donovan, M., 2012. Genetic architectures of psychiatric disorders: the emerging picture and its implications. Nat. Rev. Genet. 13, 537–551.
Sullivan, P.F., Neale, M.C., Kendler, K.S., 2000. Genetic epidemiology of major depression: review and meta-analysis. Am. J. Psychiatry 157, 1552–1562.
The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans, 2015. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 348, 648–660.
The Genotype-Tissue Expression (GTEx) project, 2013. The genotype-tissue expression (GTEx) project. Nat. Genet. 45, 580–585.
Tobe, E.H., 2013. Mitochondrial dysfunction, oxidative stress, and major depressive disorder. Neuropsychiatr. Dis. Treat. 9, 567–573.
Visconte, V., Avishai, N., Mahfouz, R., Tabarroki, A., Cowen, J., Sharghi-Moshtaghin, R., Hitomi, M., Rogers, H.J., Hasrouni, E., Phillips, J., Sekeres, M.A., Heuer, A.H., Sauntharajah, Y., Barnard, J., Tiu, R.V., 2015. Distinct iron architecture in SF3B1-mutant myelodysplastic syndrome patients is linked to an SLC25A37 splice variant with a retained intron. Leukemia 29, 188–195.
Wang, Y., Langer, N.B., Shaw, G.C., Yang, G., Li, L., Kaplan, J., Paw, B.H., Bloomer, J.R., 2011. Abnormal mitoferrin-1 expression in patients with erythropoietic protoporphyria. Exp. Hematol. 39, 784–794.
Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., Parkinson, H., 2014. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic. acids. Res. 42, D1001–D1006.
Westra, H.J., Peters, M.J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., Christiansen, M.W., Fairfax, B.P., Schramm, K., Powell, J.E., Zhernakova, A., Zhernakova, D.V., Veldink, J.H., Van den Berg, L.H., Karjalainen, J., Withoff, S., Uitterlinden, A.G., Hofman, A., Rivadeneira, F., t Hoen, P.A., Reinmaa, E., Fischer, K., Nelis, M., Milani, L., Melzer, D., Ferrucci, L., Singleton, A.B., Hernandez, D.G., Nalls, M.A., Homuth, G., Nauck, M., Radke, D., Volker, U., Perola, M., Salomaa, V., Brody, J., Suchy-Dicey, A., Gharib, S.A., Enquobahrie, D.A., Lumley, T., Montgomery, G.W., Makino, S., Prokisch, H., Herder, C., Roden, M., Grallert, H., Meitinger, T., Strauch, K., Li, Y., Jansen, R.C., Visscher, P.M., Knight, J.C., Psaty, B.M., Ripatti, S., Teumer, A., Frayling, T.M., Metspalu, A., van Meurs, J.B., Franke, L., 2014. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat. Genet. 45, 1238–1243.
Willer, C.J., Li, Y., Abecasis, G.R., 2010. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26, 2190–2191.
Wittkopp, P.J., 2005. Genomic sources of regulatory variation in cis and in trans. Cell. Mol. Life. Sci. 62, 1779–1783.
Wray, N.R., Pergadia, M.L., Blackwood, D.H., Penninx, B.W., Gordon, S.D., Nyholt, D.R., Ripke, S., MacIntyre, D.J., McGhee, K.A., Maclean, A.W., Smit, J.H., Hottenga, J.J., Willemsen, G., Middeldorp, C.M., de Geus, E.J., Lewis, C.M., McGuffin, P., Hickie, I.B., van den Oord, E.J., Liu, J.Z., Macgregor, S., McEvoy, B.P., Byrne, E.M., Medland, S.E., Statham, D.J., Henders, A.K., Heath, A.C., Montgomery, G.W., Martin, N.G., Boomsma, D.I., Madden, P.A., Sullivan, P.F., 2011. Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. Mol. Psychiatry 17, 36–48.
Wulsin, L.R., Vaillant, G.E., Wells, V.E., 1999. A systematic review of the mortality of depression. Psychosom. Med. 61, 6–17.
Yamamoto, M., Arimura, H., Fukushige, T., Minami, K., Nishizawa, Y., Tanimoto, A., Kanekura, T., Nakagawa, M., Akiyama, S., Furukawa, T., 2014. Abcb10 role in heme biosynthesis in vivo: Abcb10 knockout in mice causes anemia with protoporphyrin IX and iron accumulation. Mol. Cell. Biol. 34, 1077–1084.

Identification of *SLC25A37* as a major depressive disorder risk gene

Yong-Xia Huo[1,4], Liang Huang[2,4], Deng-Feng Zhang[1], Yong-Gang Yao[1], Yi-Ru Fang[3], Chen Zhang[3,*], Xiong-Jian Luo[1,*]

Supplementary Table 1. Demographics of MDD cases and controls for expression analyses

	Cases	Controls	*P*
Number of subjects (n)	50	50	
Age (years), mean (SD)	29.2 (6.0)	30.8 (6.1)	0.19
Gender, male n (%)	17 (34.0)	21 (42.0)	0.54
Smoking status, n (%)	13 (26.0)	15 (30.0)	0.82
Alcoholic abuse, n (%)	0 (0)	0 (0)	
Duration of illness (month) a, mean (SD)	2.9 (1.0)	N/A	
HRSD-17 b, mean (SD)	24.9 (2.2)	N/A	
Number of episode, mean (SD)	1.4 (0.4)	N/A	
Family history of mood disorders, n (%)	3 (6.0)	N/A	

Note: [a]Duration of illness prior to admission
[b]HRSD-17 on admission

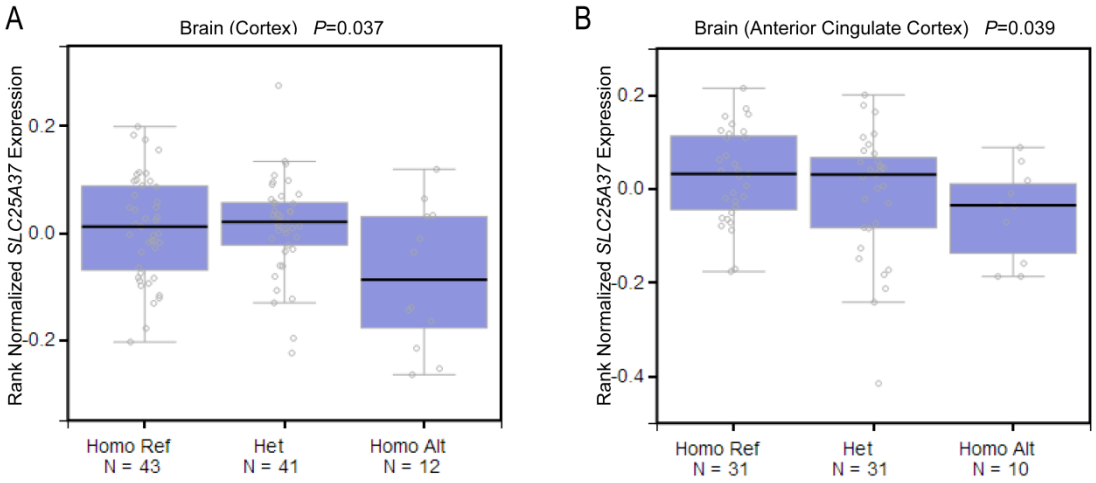

Figure S1. SNP rs6983724 is significantly associated with SLC25A37 expression in brain tissues (A, Cortex; B, Anterior Cingulate Cortex). Expression data was extracted from The GTEx consortium (1).

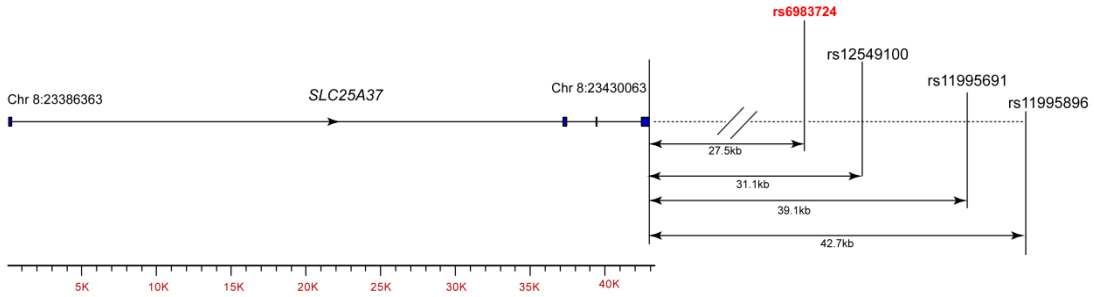

Figure S2. Genomic structure of *SLC25A37* gene and SNP rs6983724.

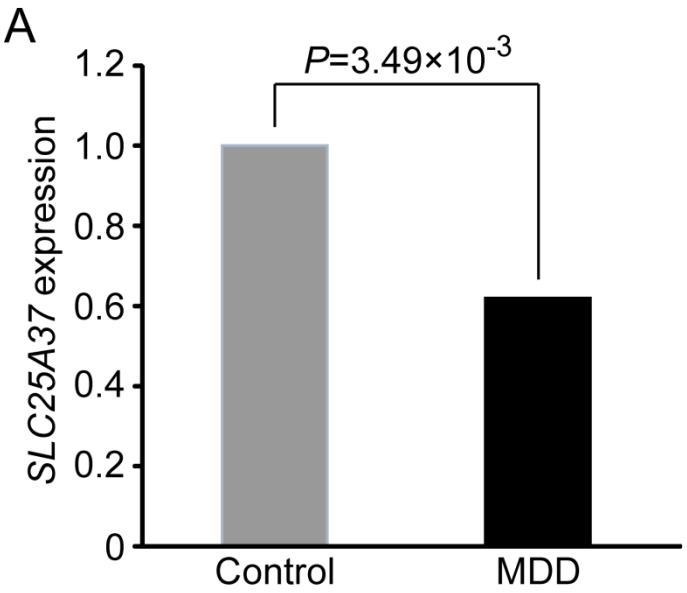

Figure S3. *SLC25A37* is significantly down-regulated in hippocampus of MDD patients. Expression data was extracted from Duric et al (2).

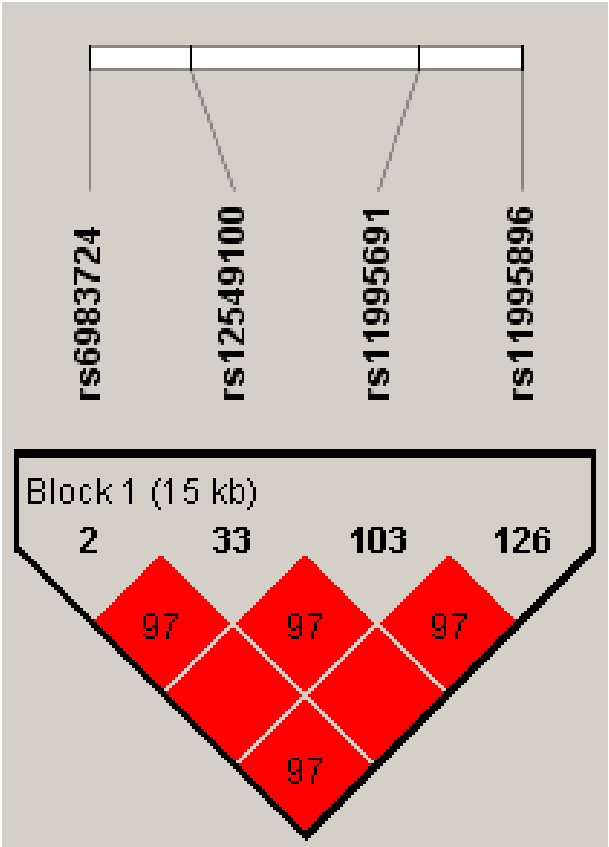

Figure S4. Linkage disequilibrium (LD) between the top four SNPs located on 8q21.1. Genotype data of European population was obtained from the 1000 genomes project (3). LD values (r^2) are shown.

References

1. The Genotype-Tissue Expression (GTEx) project (2013): The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 45:580-585.
2. Duric V, Banasr M, Licznerski P, Schmidt HD, Stockmeier CA, Simen AA, et al. (2010): A negative regulator of MAP kinase causes depressive behavior. *Nat Med*. 16:1328-1332.
3. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, et al. (2010): A map of human genome variation from population-scale sequencing. *Nature*. 467:1061-1073.