

Loci with genome-wide associations with schizophrenia in the Han Chinese population

Zhiqiang Li,* Yuqian Xiang,* Jianhua Chen, Qiaoli Li, Jiawei Shen, Yun Liu, Wenjin Li, Qinghe Xing, Qingzhong Wang, Lei Wang, Guoyin Feng, Lin He, Xinzhi Zhao and Yongyong Shi

Background

A large schizophrenia genome-wide association study (GWAS) and a subsequent extensive replication study of individuals of European ancestry identified eight new loci with genome-wide significance and suggested that the *MIR137*-mediated pathway plays a role in the predisposition for schizophrenia.

Aims

To validate the above findings in a Han Chinese population.

Method

We analysed the single nucleotide polymorphisms (SNPs) in the newly identified schizophrenia candidate loci and predicted *MIR137* target genes based on our published Han Chinese populations (BIOX) GWAS data. We then analysed 18 SNPs from the candidate regions in an independent cohort that consisted of 3585 patients with schizophrenia and 5496 controls of Han Chinese ancestry.

Results

We replicated the associations of five markers ($P < 0.05$),

including three that were located in the predicted *MIR137* target genes. Two loci (*ITIH3/4*: rs2239547, $P = 1.17 \times 10^{-10}$ and *CALN1*: rs2944829, $P = 9.97 \times 10^{-9}$) exhibited genome-wide significance in the Han Chinese population.

Conclusions

The *ITIH3/4* locus has been reported to be of genome-wide significance in the European population. The successful replication of this finding in a different ethnic group provides stronger evidence for the association between schizophrenia and *ITIH3/4*. We detected the first genome-wide significant association of schizophrenia with *CALN1*, which is a predicted target of *MIR137*, and thus provide new evidence for the associations between *MIR137* targets and schizophrenia.

Declaration of interest

None.

Copyright and usage

© The Royal College of Psychiatrists 2015.

Schizophrenia is a chronic, severe and disabling psychiatric disorder that affects 1% of the population worldwide. Schizophrenia is a complex major disease that manifests as psychotic behaviour (delusions and hallucinations), disorganisation, dysfunction in normal affective responses and altered cognitive functioning.¹ Twin, family and adoption studies have demonstrated that schizophrenia is highly heritable; the heritability of schizophrenia is estimated to be as high as 80%.^{2,3} However, the genetic loci that contribute to the disease remain generally elusive. Recent genome-wide association studies (GWAS) have provided unbiased assessments of common sequence variations across the whole genome and may robustly map the loci involved in the pathology of complex diseases. In the past 5 years, a number of GWAS of schizophrenia have been published.^{4–20} Several loci have surpassed the genome-wide significance threshold ($P = 5 \times 10^{-8}$) in more than one GWAS; for example the major histocompatibility complex (MHC) region has exhibited genome-wide significance in four independent large-scale GWAS of schizophrenia.^{9,14,20,21}

Recently, the Schizophrenia Psychiatric GWAS Consortium (PGC) reported a large schizophrenia GWAS.¹³ In this study, the authors conducted a mega-analysis of the combined genotyping data from 21 856 individuals of European ancestry from 17 separate studies and then performed a replication study in a sample of 29 839 participants from 19 populations. This study revealed seven loci with genome-wide significance, including two previously reported loci (i.e. the MHC region and *TCF4*) and the following five novel loci: 1p23.3 (*MIR137*), 2q32.3 (*PCGEM1*), 8p23.2 (*CSMD1*), 8q21.3 (*MMP16*), and 10q24.32–q24.33 (*CNNM2/NT5C2*). Moreover, a joint analysis of schizophrenia and bipolar disorder identified three additional genes that reached genome-wide significance:

CACNA1C, *ANK3* and *ITIH3/4*. The most significant new finding was the identification of *MIR137*, which encodes the microRNA 137, which is a known regulator of neuronal development. Notably, microRNA 137 may directly regulate some other schizophrenia susceptibility genes. Among the 301 high-confidence predicted *MIR137* targets, 17 had at least one significant single nucleotide polymorphism (SNP) at $P < 10^{-4}$; these 17 targets included four genome-wide significant genes (i.e. *TCF4*, *CACNA1C*, *CSMD1* and *C10orf26*). Subsequently, these four genes and *ZNF804A*, another compelling candidate gene for schizophrenia, were validated as *MIR137* targets.^{22,23} These findings suggest that the *MIR137*-mediated pathway is involved in the aetiology of schizophrenia. In a replication study, Hamshere *et al* tested 78 of the 81 SNPs highlighted by the PGC in a UK population and found significant association for 37 (47%) of the SNPs. Remarkably, genetic variants in three new loci (i.e. *ITIH3/4*, *CACNA1C* and *SDCCAG8*) reached genome-wide significance after combining their new schizophrenia data (CLOZUK) with those of the PGC.²⁴ The CLOZUK sample is a series of 2640 UK individuals that were registered for clozapine treatment and had clinical diagnoses of schizophrenia and 2878 controls.²⁴

Only a few schizophrenia GWAS have been conducted in non-Western populations. One of these was our GWAS of the genotypes of 3750 individuals with schizophrenia and 6468 healthy controls from Han Chinese populations (BIOX GWAS).¹⁵ None of the SNPs within the 10 loci reported by PGC met the criteria for genome-wide significance in our data-set. However, the criteria for genome-wide significance minimises the occurrence of false positives (i.e. type I errors), whereas the occurrence of false negatives (i.e. type II errors) was almost certain. In another schizophrenia GWAS that was conducted in a Han Chinese

*These authors contributed equally to the work.

population, Yue *et al*¹⁴ identified the susceptibility locus at the MHC region. We ascertained significant associations between the MHC region, the *TCF4* gene and schizophrenia in our previous study of 2496 patients with schizophrenia and 5184 normal controls drawn from a Han Chinese population, but we failed to confirm an association between the *NRGN* gene and schizophrenia.²⁵ Regarding the psychiatric traits, most of the associations were population specific; however, in some cases, the associations may exhibit convergence of risk genes (but not necessarily risk alleles) across populations.²⁶ It is of great interest to replicate the findings related to the other loci that were identified by the PGC in the Han Chinese population. In the present study, we conducted a two-stage analysis. We first analysed the SNPs ($n = 2595$) of the eight newly identified loci, predicted the targets of microRNA 137 in the BIOX GWAS data and selected 18 candidate SNPs. These SNPs were genotyped and tested for their associations with schizophrenia in a replication cohort that consisted of 3585 patients with schizophrenia (the schizophrenia group) and 5496 controls (the control group) of Han Chinese ancestry. A meta-analysis was performed to combine the Chinese data-sets (BIOX GWAS and replication).

Method

Participants

Participants in the schizophrenia group were in-patients or out-patients who were recruited from various mental health centres. The patients were interviewed by two independent psychiatrists and were diagnosed according to DSM-IV criteria,²⁷ and had 2-year histories of the disorder. All met the following two criteria: preoccupation with one or more delusions and frequent auditory hallucinations. However, none of the following symptoms were prominent: disorganised speech, disorganised or catatonic behaviour, or flat or inappropriate affect. All healthy controls were randomly selected from Chinese Han volunteers (from hospitals and a community survey) who were asked to reply to a written invitation to evaluate their medical histories. Potential lists of controls were screened for suitable volunteers by excluding individuals with major mental illnesses. The sample consisted of 3585 people in the schizophrenia group (1901 men and 1684 women, the mean age at onset of schizophrenia was 35.0 years, s.d. = 11.0) and 5496 controls (2819 men and 2677 women with a mean age of 46.4 years, s.d. = 14.1). Of the participants, 1329 in the schizophrenia group and 2037 in the control group were northern Han Chinese; 1700 in the schizophrenia group and 2622 in the control group were central Han Chinese; and 556 in the schizophrenia group and 837 in the control group were southern Han Chinese. Approval was received for our study from the local Human Genetic Resources Ethics Committee of Human Genetic Resources. After providing a complete description of the study to the participants, written informed consent was obtained.

BIOX GWAS quality control

The gender established via data genotyping was checked for each of the participants, and individuals in the schizophrenia group of unknown or inconsistent gender (compared with the sample record) were removed ($n = 49$). Arrays with call rates $< 95\%$ were excluded ($n = 276$). SNPs with call rates $< 95\%$ in either the schizophrenia or the control group were removed ($n = 92\,324$). SNPs with minor allele frequencies $< 3\%$ ($n = 228\,267$) and those that significantly deviated from Hardy–Weinberg Equilibrium (HWE; $P \leq 1 \times 10^{-6}$) among the control group ($n = 28\,657$) were also excluded. Heterozygosity rates were calculated with the intent of removing deviations that exceeded six standard deviations

from the mean; however, no samples were excluded based on this criterion. PLINK's identity by descent analysis was used to detect cryptic relatedness. When a pair of individuals exhibited a $PI_HAT > 0.25$, the member of the pair with the lower call rate was excluded from the analysis ($n = 146$). Population ancestry assessments were evaluated using principal components analysis, and all of the samples were of Chinese ancestry.

SNP selection and genotyping

The allelic frequencies of the nine genome-wide significant SNPs reported in the PGC and PGC+CLOZUK studies differ widely between the European and Chinese populations (see online Table DS1, which is based on HAPMAP CEU (Utah residents with Northern and Western European ancestry from the Centre d'Etude du Polymorphisme Humain project collection) and CHB (Han Chinese in Beijing, China) data-sets), which indicates genetic heterogeneity. Therefore, we tended to select the SNPs that exhibited more significant associations in the BIOX GWAS dataset. Moreover, the linkage disequilibrium information about the chosen SNPs and those reported in the PGC and PGC+CLOZUK studies was also considered. Ultimately, 18 SNPs from 12 genes were selected for this study (online Table DS2) and further information about the selection of SNPs is provided in the Results section.

The MassARRAY iPLEX Gold platform (Sequenom, San Diego, California, USA) was used for genotyping. Polymerase chain reaction (PCR) amplification primers and single-base extension unextended primers were designed using the online software Human Genotyping Tools (www.mysequenom.com/Tools). The genotyping analysis was performed according to the manufacturer's protocol (Sequenom). Genotype callings were extracted by the MassARRAY Typer software (version 4.0). In total 9173 individuals, including 3648 in the schizophrenia group and 5525 control groups, were genotyped. Individuals with less than 90% call rates (which may indicate poor DNA quality) were excluded from the analysis. Therefore, 9081 samples (3585 in the schizophrenia group and 5496 in the control group) were kept for further analysis. For each SNP, the call rates and P -values that resulted from HWE testing in the control group were calculated for data quality control.

Analyses

HapMap genotypes and haplotype data were downloaded using Haploview.²⁸ HWE tests, allelic association tests and meta-analysis of each SNP were conducted using PLINK.²⁹ The heterogeneity across studies was evaluated using the Q -test. The Mantel–Haenszel method was used to calculate the fixed-effect estimate. We stratified our samples into northern, central and southern groups according to their geographic region, performed analysis of each group and then combined the results via meta-analysis in the replication study.

Results

Comparison of the BIOX and PGC/PGC+CLOZUK GWAS data-sets in terms of the newly identified loci

Nine genome-wide significant SNPs within the eight loci were reported in the PGC and PGC+CLOZUK studies (PGC: 1p21.3, 2q32.3, 8p23.2, 8q21.3 and 10q24.32–q24.33, PGC+CLOZUK: *ITIH3/4*, *CACNA1C* and *SDCCAG8*), and their frequencies differed widely between the European and Chinese populations (Table DS1). In the PGC study, the strongest association signal was observed in locus 1p21.3 (*MIR137*, rs1625579, $P = 1.59 \times 10^{-11}$).

rs1625529 was not genotyped in the BIOX GWAS data-set; however, a proxy SNP (rs1198588, $D' = 1$ and $r^2 = 1$ in CHB) was present. This proxy SNP exhibited marginal significance ($P = 6.91 \times 10^{-2}$), and the direction of the effect size was consistent with the PGC report. Therefore, the rs1198588 SNP was selected in the follow-up stage. In 10q24.33 (*NT5C2*), the reported genome-wide significant SNP was rs11191580 ($P = 1.11 \times 10^{-8}$) in the PCG study. In the follow-up phase, we selected rs732998 ($P = 9.15 \times 10^{-3}$ in the BIOX GWAS data-set), which is in complete linkage disequilibrium with rs11191580 ($D' = 1$, $r^2 = 1$) in the HAPMAP CHB data-set. In the locus *ITIH3/4*, no significant SNPs were observed in the BIOX GWAS data-set. However, rs2239547 was not genotyped in the BIOX GWAS, but the SNP was included in the follow-up phase. Within the other five loci (i.e. 8p23.2, 8q21.3, 10q24.32, *SDCCAG8* and *CACNA1C*), the genome-wide significant SNPs reported in the PGC or PGC+CLOZUK studies were not genotyped or did not exhibit associations in the BIOX GWAS data-set; thus, we selected the other SNPs within each locus that exhibited the most significant associations ($P < 0.05$) for replacement in the follow-up study. In the loci with multiple markers that were found to be significant in the BIOX GWAS data-set, all of the SNPs with P -values below 0.01 were selected. For each locus, at least one SNP was selected with the exception of locus 2q32.3 because no significant SNPs were observed in this locus in the BIOX GWAS data-set and because, as reported in the PGC study, rs17662626 was non-polymorphic in the HAPMAP CHB population.

Results for the 17 top-scoring *MIR137* predicted target genes in the BIOX GWAS data-set

Among the 17 *MIR137* predicted target genes (i.e., *C10orf26*, *TCF4*, *CSMD1*, *CACNA1C*, *SLC12A2*, *CALN1*, *GRIA1*, *ST13*, *CSDC2*, *LUZP2*, *GLIS2*, *EPHA7*, *CADPS2*, *RGS6*, *TBC1D12*, *FAM78A* and *C20orf108*), a total of 15 (i.e. all but *CSDC2* and *GLIS2*) with at least one SNP were genotyped in the BIOX GWAS data-set. Within these regions, 2230 genotyped SNPs were analysed and 29 SNPs (1.30%) at *CSMD*, *CACNA1C*, *CALN1*, *CADPS2* and *RGS6* exhibited significance at the level of $P < 0.01$. Ten tagSNPs were selected for the follow-up phase.

Results of the follow-up phase and combined analysis

The detailed SNP quality control information is given in Table DS2, and the full results for all SNPs are listed in online Table DS3. In the follow-up study, we replicated the associations of five markers ($P < 0.05$, with effects in the same direction, and P -values of less than 0.001 in the meta-analysis of the combined Han Chinese samples). The summary results for the five SNPs are shown in Table 1. The SNP with the most significant association was located in an intron of *ITIH3/4* (rs2239547, odds ratio (OR) = 0.81, $P = 1.17 \times 10^{-10}$), which is consistent with Hamshere *et al*'s report²⁴ (OR = 0.90, $P = 3.62 \times 10^{-10}$ in the PGC+CLOZUK data-set). Meta-analysis of the combined European and Chinese samples ($P = 7.54 \times 10^{-17}$, OR = 0.88; the Cochran's Q -test of heterogeneity $P > 0.01$) strongly supported the association of this SNP.

The second strongest association was observed in *CALN1* (rs2944829, OR = 0.85, $P = 3.17 \times 10^{-6}$). In the combined BIOX GWAS and follow-up analysis, this locus reached genome-wide significance ($P = 9.97 \times 10^{-9}$, (A allele) OR = 0.85, 95% CI 0.81–0.90; no significant heterogeneity was observed). The PGC reported that rs12699131 (35.4 kb upstream of rs2944829 (A allele) OR = 1.10, $P = 5.16 \times 10^{-6}$) was associated with schizophrenia in

Table 1 Summary of the Han Chinese population (BIOX) genome-wide association study (GWAS), replication and meta-analyses for the top five most significant single nucleotide polymorphisms (SNPs)^a

Chromosome	SNP	Position	A ^b	BIOX GWAS			Replication			Meta-analysis			Gene
				P	OR for A1 (95% CI)	P	OR for A1 (95% CI)	P	OR for A1 (95% CI)	P	OR for A1 (95% CI)		
3	rs2239547	52855229	G	–	–	1.17 × 10 ⁻¹⁰	0.81 (0.76–0.87)	–	–	–	–	–	<i>ITIH3/4</i>
7	rs2944829	71424657	A	8.34 × 10 ⁻⁰⁴	0.86 (0.79–0.94)	3.17 × 10 ⁻⁰⁶	0.85 (0.80–0.91)	9.97 × 10 ⁻⁰⁹	0.85 (0.81–0.90)	–	–	–	<i>CALN1</i>
7	rs2192017	122054147	A	1.20 × 10 ⁻⁰³	1.13 (1.05–1.22)	2.78 × 10 ⁻⁰²	1.08 (1.01–1.15)	1.61 × 10 ⁻⁰⁴	1.10 (1.05–1.15)	–	–	–	<i>CADPS2</i>
10	rs10748844	105315160	T	3.94 × 10 ⁻⁰³	1.12 (1.04–1.22)	1.92 × 10 ⁻⁰²	1.08 (1.01–1.15)	2.96 × 10 ⁻⁰⁴	1.09 (1.04–1.15)	–	–	–	<i>NEURL</i>
12	rs2887780	2369887	C	6.11 × 10 ⁻⁰³	0.90 (0.83–0.97)	1.21 × 10 ⁻⁰³	0.90 (0.85–0.96)	2.25 × 10 ⁻⁰⁵	0.90 (0.86–0.95)	–	–	–	<i>CACNA1C</i>

a. The position is based on the National Center for Biotechnology Information (NCBI) Genome browser build 36.1. rs2239547 was not available in our GWAS sample.
b. A1, minor allele name (based on the whole sample).

the European-ancestry populations.¹³ The HapMap data-set revealed that rs12699131 was in tight linkage disequilibrium with rs2944829 in the CEU+TSI (Toscani in Italia) populations ($D' = 0.94$, $r^2 = 0.68$) and that the frequencies of the rs12699131 and rs2944829 haplotypes of the risk allele (A–G) and the protective allele (G–A) were 38.8% and 52.2% respectively. These results suggested that our finding of an association (rs2944829) is consistent with the findings (rs12699131) from individuals of European descent.

Of the eight genome-wide significant schizophrenia candidate loci identified in the European populations, the associations at *ITIH3/4* (genome-wide significance) and *CACNA1C* ($P < 0.001$) were replicated in our study. At locus 1p21.3, rs1198588 (*MIR137*) did not exhibit a significant association in the follow-up study. However, variants in three *MIR137* target genes (i.e., *CALN1*, *CADPS2* and *CACNA1C*) were associated with schizophrenia in the Han Chinese population at $P < 0.001$, and the variant in *CALN1* attained genome-wide significance.

Discussion

Main findings

To validate the findings of the PGC and PGC+CLOZUK schizophrenia studies (i.e. the European GWAS data-sets) in the Han Chinese population, we selected 18 SNPs from 12 schizophrenia candidate genes based on the BIOX GWAS data and then performed an association study in an extended Han Chinese replication cohort. We replicated the associations of 5/18 markers at nominal thresholds ($P < 0.05$) with $P < 0.001$ in the combined analysis of the BIOX GWAS and the replication, and two markers (rs2239547 at *ITIH3/4*, $P = 1.17 \times 10^{-10}$ and rs2944829 at *CALN1*, $P = 9.97 \times 10^{-9}$) reached genome-wide significance. Both of our association signals are consistent with the European findings.

Significance of our findings

Rs2239547 occurred in an extensive linkage disequilibrium block that contained many genes.²⁴ However, rs2239547 is an intronic SNP of *ITIH4*, and the expression quantitative trait loci (eQTL) data-set showed that it is significantly associated with the expression of *ITIH4* ($P = 3.51 \times 10^{-11}$) in lymphoblastoid cell lines from HapMap JPT (Japanese in Tokyo, Japan) and CHB populations.³⁰ Moreover, the *ITIH4* protein has been found to be completely absent in patients who have had acute ischaemic strokes.³¹ Moreover, patients with schizophrenia have a greater likelihood of developing stroke than do controls.^{32,33} Thus, for this locus, *ITIH4* is one of the most compelling functional candidates for further study.

We detected the first genome-wide significance of *CALN1* in schizophrenia via joint analysis of the BIOX GWAS data. *CALN1* has two conserved EF-hand-type calcium-binding motifs and is exclusively and highly expressed in the brain, which suggests that this gene has a role in calcium signalling in the central nervous system.³⁴ Interestingly, a recent genome-wide association analysis that identified 13 new risk loci for schizophrenia suggested involvement of neuronal calcium signalling in the aetiology of schizophrenia.³⁵ Additionally, several patients with *CALN1* deletions have been found to exhibit intellectual disabilities.³⁶ Cognitive impairment is an important clinical feature of schizophrenia. We suggest that further research may clarify whether *CALN1* is associated with neurocognitive phenotypes.

MIR137 and five of its targets (*TCF4*, *CACNA1C*, *CSMD1*, *C10orf26* and *ZNF804A*) were reported to be associated with schizophrenia in the GWAS of individuals of European ancestry,

which suggests that *MIR137*-mediated dysregulation is an important aetiological mechanism of schizophrenia.¹³ In this study, we failed to directly replicate the association between *MIR137* and schizophrenia in the Han Chinese. However, genetic variants in three predicted *MIR137*-target genes (i.e. *CALN1*, *CADPS2* and *CACNA1C*) were significantly associated with schizophrenia; these variants included rs2944829, which reached genome-wide significance. Our findings support the hypothesis that the *MIR137* pathway is involved in the aetiology of schizophrenia.

Within the eight genome-wide significant loci that were reported by the PGC or PGC+CLOZUK studies, the genome-wide significance of one locus (*ITIH3/4*) was replicated in our study, which provides evidence that overlapping polygenic variation exists between ethnically divergent populations. However, the other loci did not reach genome-wide significance. One possible reason for the non-replication of some genes is that the selected proxy SNPs are in extremely low linkage disequilibrium with the original genome-wide significant SNPs. These failures may also partially be explained by genetic heterogeneity across the different populations. However, other factors, such as polygenic heterogeneity, should also be considered. Moreover, the strongest signal (*MIR137* locus) reported in the PGC study also failed to be replicated in the CLOZUK study. All of these phenomena highlight the genetic complexity of schizophrenia, and some susceptibility for schizophrenia is likely to result from population-specific variants. Meta-analyses and mega-analyses across different populations and larger sample sizes would efficiently aid the unravelling of this complexity. In this study, we enlarged the sample size to validate the European findings.

In summary, we independently confirmed the association of schizophrenia with *ITIH3/4*, which was previously reported to have genome-wide significance, and identified the genome-wide significance of *CALN1*, which had not previously attained genome-wide significance. Replication across different ethnic groups provides stronger evidence for the associations between schizophrenia and these two loci, and their biological mechanisms will become increasingly important for understanding of the aetiology of schizophrenia.

Zhiqiang Li, PhD, Children's Hospital, Fudan University, Shanghai and Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai; **Yuqian Xiang**, PhD, Children's Hospital, Fudan University, Shanghai; **Jianhua Chen**, PhD, Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai; **Qiaoli Li**, PhD, Children's Hospital and Institutes of Biomedical Sciences, Fudan University, Shanghai; **Jiawei Shen**, PhD, Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai; **Yun Liu**, PhD, Institutes of Biomedical Sciences, Fudan University, Shanghai; **Wenjin Li**, PhD, Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai; **Qinghe Xing**, PhD, Institutes of Biomedical Sciences, Fudan University, Shanghai; **Qingzhong Wang**, PhD, Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai; **Lei Wang**, PhD, Institutes of Biomedical Sciences, Fudan University, Shanghai; **Guoyin Feng**, Shanghai Institute of Mental Health, Shanghai; **Lin He**, PhD, Children's Hospital, Fudan University, Shanghai and Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai; **Xinzhong Zhao**, PhD, Children's Hospital, Fudan University, Shanghai; **Yongyong Shi**, PhD, Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, Shanghai, People's Republic of China

Correspondence: Yongyong Shi, Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai Jiao Tong University, 1954 Huashan Road, Shanghai 200030, People's Republic of China. Email: shiyongyong@gmail.com; Xinzhong Zhao, Children's Hospital, Fudan University, 138 Yixueyuan Road, Shanghai 200032, People's Republic of China. Email: xzzhou@fudan.edu.cn

First received 23 Apr 2014, final revision 7 Nov 2014, accepted 11 Dec 2014

Funding

Design and conduct of the study: collection, management, analysis and interpretation of the data: the Natural Science Foundation, Ministry of Science and Technology and the Ministry of Education of China. Preparation, review or approval of the manuscript: Shanghai Municipal Science & Technology Commission, Education Commission and Education Development Foundation.

References

- van Os J, Kapur S. Schizophrenia. *Lancet* 2009; **374**: 635–45.
- Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* 2009; **373**: 234–9.
- Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry* 2003; **60**: 1187–92.
- Lencz T, Morgan TV, Athanasiou M, Dain B, Reed CR, Kane JM, et al. Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. *Mol Psychiatry* 2007; **12**: 572–80.
- Shifman S, Johannesson M, Bronstein M, Chen SX, Collier DA, Craddock NJ, et al. Genome-wide association identifies a common variant in the reelin gene that increases the risk of schizophrenia only in women. *PLoS Genet* 2008; **4**: e28.
- Kirov G, Zaharieva I, Georgieva L, Moskvina V, Nikolov I, Cichon S, et al. A genome-wide association study in 574 schizophrenia trios using DNA pooling. *Mol Psychiatry* 2008; **14**: 796–803.
- Sullivan PF, Lin D, Tzeng JY, van den Oord E, Perkins D, Stroup TS, et al. Genomewide association for schizophrenia in the CATIE study: results of stage 1. *Mol Psychiatry* 2008; **13**: 570–84.
- Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Sci Signalling* 2008; **320**: 539.
- Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D, et al. Common variants conferring risk of schizophrenia. *Nature* 2009; **460**: 744–7.
- Need AC, Ge D, Weale ME, Maia J, Feng S, Heinzen EL, et al. A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet* 2009; **5**: e1000373.
- O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V, et al. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* 2008; **40**: 1053–5.
- Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–52.
- Schizophrenia Psychiatric Genome-Wide Association Study C. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–76.
- Yue WH, Wang HF, Sun LD, Tang FL, Liu ZH, Zhang HX, et al. Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2. *Nat Genet* 2011; **43**: 1228–31.
- Shi Y, Li Z, Xu Q, Wang T, Li T, Shen J, et al. Common variants on 8p12 and 1q24.2 confer risk of schizophrenia. *Nat Genet* 2011; **43**: 1224–7.
- Bergen SE, O'Dushlaine CT, Ripke S, Lee PH, Ruderfer DM, Akterin S, et al. Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. *Mol Psychiatry* 2012; **17**: 880–6.
- Strange A, Riley BP, Spencer CCA, Morris DW, Pirinen M, O'Dushlaine CT, et al. Genome-wide association study implicates HLA-C* 01: 02 as a risk factor at the major histocompatibility complex locus in schizophrenia. *Biol Psychiatry* 2012; **72**: 620–8.
- Levinson DF, Shi J, Wang K, Oh S, Riley B, Pulver AE, et al. Genome-wide association study of multiplex schizophrenia pedigrees. *Am J Psychiatry* 2012; **169**: 963–73.
- Betcheva ET, Yosifova AG, Mushiroda T, Kubo M, Takahashi A, Karachanak SK, et al. Whole-genome-wide association study in the Bulgarian population reveals HHAT as schizophrenia susceptibility gene. *Psychiatr Genet* 2013; **23**: 11–9.
- Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, et al. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009; **460**: 753–7.
- Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–52.
- Kim AH, Parker EK, Williamson V, McMichael GO, Fanous AH, Vladimirov VI. Experimental validation of candidate schizophrenia gene ZNF804A as target for hsa-miR-137. *Schizophrenia Res* 2012; **141**: 60–4.
- Kwon E, Wang W, Tsai LH. Validation of schizophrenia-associated genes CSMD1, C10orf26, CACNA1C and TCF4 as miR-137 targets. *Mol Psychiatry* 2013; **18**: 11–2.
- Hamshere ML, Walters JT, Smith R, Richards AL, Green E, Grozeva D, et al. Genome-wide significant associations in schizophrenia to ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by the Schizophrenia PGC. *Mol Psychiatry* 2013; **18**: 708–12.
- Li T, Li Z, Chen P, Zhao Q, Wang T, Huang K, et al. Common variants in major histocompatibility complex region and TCF4 gene are significantly associated with schizophrenia in Han Chinese. *Biol Psychiatry* 2010; **68**: 671–3.
- Gelernter J, Kranzler HR, Sherva R, Almasy L, Koesterer R, Smith AH, et al. Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci. *Mol Psychiatry* 2014; **19**: 41–9.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (4th edn) (DSM-IV)*. APA, 1994.
- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–75.
- Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, et al. Population genomics of human gene expression. *Nat Genet* 2007; **39**: 1217–24.
- Kashyap RS, Nayak AR, Deshpande PS, Kabra D, Purohit HJ, Taori GM, et al. Inter-alpha-trypsin inhibitor heavy chain 4 is a novel marker of acute ischemic stroke. *Clinica Chimica Acta* 2009; **402**: 160–3.
- Tsai KY, Lee CC, Chou YM, Su CY, Chou FHC. The incidence and relative risk of stroke in patients with schizophrenia: a five-year follow-up study. *Schizophr Res* 2012; **138**: 41–7.
- Lin HC, Hsiao FH, Pfeiffer S, Hwang YT, Lee HC. An increased risk of stroke among young schizophrenia patients. *Schizophr Res* 2008; **101**: 234–41.
- Wu YQ, Lin X, Liu CM, Jamrich M, Shaffer LG. Identification of a human brain-specific gene, calneuron 1, a new member of the calmodulin superfamily. *Mol Genet Metab* 2001; **72**: 343–50.
- Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet* 2013; **45**: 1150–9.
- Slavotinek AM, Rosenfeld JA, Chao R, Niyazov D, Eswara M, Bader PI, et al. A de novo deletion of CALN1 in a male with a bilateral diaphragmatic defect does not definitely cause this malformation. *Am J Med Genet A* 2011; **155**: 1196–201.

