Common variants of *IRF3* conferring risk of schizophrenia

Xiao Li [a, c, 1], Wen Zhang [a, 1], Todd Lencz [d, e], Ariel Darvasi [f], Anna Alkelai [g], Bernard Lerer [g], Hong-Yan Jiang [h, i], Deng-Feng Zhang [a, c], Li Yu [h], Xiu-feng Xu [i], Ming Li [b, *], Yong-Gang Yao [a, c, j, **]

[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, China
[b] Lieber Institute for Brain Development, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[c] Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan, China
[d] The Zucker Hillside Hospital, Psychiatry Research, 75-59 263rd Street, Glen Oaks, NY, USA
[e] Feinstein Institute for Medical Research, 350 Community Drive Manhasset, NY, USA
[f] Department of Genetics, Institute of Life Sciences, The Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel
[g] Biological Psychiatry Laboratory, Department of Psychiatry, Hadassah - Hebrew University Medical Center, Jerusalem, Israel
[h] Laboratory for Conservation and Utilization of Bio-resource & Key Laboratory for Microbial Resources of the Ministry of Education, Yunnan University, Kunming, Yunnan, China
[i] Department of Psychiatry, the First Affiliated Hospital of Kunming Medical University, Kunming, Yunnan, China
[j] CAS Center for Excellence in Brain Science, Chinese Academy of Sciences, Shanghai, China

ARTICLE INFO

ABSTRACT

Schizophrenia is a brain disorder with high heritability. Recent studies have implicated genes involved in the immune response pathway in the pathogenesis of schizophrenia. Interferon regulatory factor 3 (*IRF3*), a virus-immune-related gene, activates the transcription of several interferon-induced genes, and functionally interacts with several schizophrenia susceptibility genes. To test whether *IRF3* is a schizophrenia susceptibility gene, we analyzed the associations of its SNPs with schizophrenia in independent population samples as well as reported data from expression quantitative trait loci (eQTL) in healthy individuals. We observed multiple independent SNPs in *IRF3* showing nominally significant associations with schizophrenia ($P < 0.05$); more intriguingly, a SNP (rs11880923), which is significantly correlated with *IRF3* expression in independent samples ($P < 0.05$), is also consistently associated with schizophrenia across different cohorts and in combined samples (odds ratio = 1.075, $P_{meta} = 2.08 \times 10^{-5}$), especially in Caucasians (odds ratio = 1.078, $P_{meta} = 2.46 \times 10^{-5}$). These results suggested that *IRF3* is likely a risk gene for schizophrenia, at least in Caucasians. Although the clinical associations of *IRF3* with diagnosis did not achieve genome-wide level of statistical significance, the observed odds ratio is comparable with other susceptibility loci identified through large-scale genetic association studies on schizophrenia, which could be regarded simply as small but detectable effects.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Schizophrenia is one of the most severe psychiatric disorders with worldwide lifetime prevalence approaching 1%, and characterized by psychotic features (delusions and hallucinations), disorganization, dysfunction in normal affective responses, and altered cognitive functions (Andreasen, 1995). Previous studies have implicated schizophrenia as an illness involved by interactions of one or more environmental insults with predisposing genetic susceptibility (Cannon et al., 2003; Caspi and Moffitt, 2006; Clarke et al., 2009). Among these environmental hazards, viral infection is one of the most widely accepted factors that could increase risk of future development of schizophrenia.

Viral infections produce considerable gene expression changes as they trigger immune defenses through type I interferons (IFNs) and the mobilization of transcription factors of the signal transducer and activator of transcription (STAT) and interferon regulatory factor (IRF) families (de Veer et al., 2001; Stark et al., 1998).

* Corresponding author.
** Corresponding author. Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China. Tel./fax: +86 871 65180085.
E-mail addresses: limingkiz@gmail.com (M. Li), yaoyg@mail.kiz.ac.cn (Y.-G. Yao).
[1] These authors contributed equally to this study.

Altered expressions of IRF genes in brains have been implicated to contribute to disrupted brain circuit development, maturation and function and result in behavioral deficits that overlap with those seen in schizophrenia and major depression (Hurlock, 2001; Schaefer et al., 2002a, 2002b). However, sensitivity to environmental stressors like viral infections shows substantial inter-individual variation, and at least part of this variation may be genetically determined and/or involved with gene−environment interactions.

The interferon regulatory factor 3 gene (*IRF3*), located on chromosome 19q13, a genomic region possibly harbors risk genes for psychiatric disorders, i.e., *APOE*, a risk gene for schizophrenia and Alzheimer's disease (Harold et al., 2009; Lambert et al., 2009; Liu et al., 2003; Seshadri et al., 2010). *IRF3* plays an important role in the innate immune system's response to viral infection (Collins et al., 2004), and the protein encoded by *IRF3* is found in an inactive cytoplasmic form that upon serine/threonine phosphorylation forms a complex with *CREBBP*, which can translocate to the nucleus and activate transcription of interferons alpha and beta, as well as other interferon-induced genes (Juang et al., 1998; Lee et al., 2014; Prinarakis et al., 2008).

Thus, *IRF3* is likely a susceptibility gene for schizophrenia based on these lines of epidemiologic and etiologic evidence. Interestingly, a known and predicted protein−protein interactions database (STRING, http://string-db.org/) showed that *IRF3* interacts with many other schizophrenia susceptibility genes (Franceschini et al., 2013), such as *AKT1*, *CREB1*, *ESR1*, and *TP53* (Fig. S1). This observation encouraged our speculation because of recent evidence about protein−protein interactions among schizophrenia risk genes (Luo et al., 2014a; Yu et al., 2014).

Here, we attempt to characterize the genetic contributions of common variants within *IRF3* to schizophrenia susceptibility in independent samples as well as to explore the potential effects on gene expression. The discovery stage involved data from a large screening schizophrenia case−control sample and two independent cohorts containing expression quantitative trait loci (eQTL) data; the replication step includes a variety of independent replications on both clinical and eQTL associations. Our results indicated that *IRF3* is likely a schizophrenia susceptibility gene.

2. Materials and methods

2.1. Case-control and family-based samples

We used the case−control subjects from the Schizophrenia Psychiatric Genomics Consortium (PGC) as our screening sample (Ripke et al., 2013). The schizophrenia PGC performed a large-scale meta-analysis by combining GWAS data in 13,833 schizophrenia cases and 18,310 controls (PGC1). All of the subjects were of European ancestry. Detailed information about sample description, including diagnostic assessments, genotyping, quality control and statistical analysis can be found in the original publication (Ripke et al., 2013).

For replication analysis, we recruited two independent schizophrenia case−control samples and five family-based samples from different populations. Detailed information on each sample, including diagnostic assessment, genotyping, and quality control has been reported previously (Aberg et al., 2013; Alkelai et al., 2011, 2012; Lencz et al., 2013; Zhang et al., 2014). In brief, the seven replication samples are: (1) the Jewish case−control sample: this sample contained 904 schizophrenia cases and 1640 controls (Lencz et al., 2013). (2) The Chinese case−control sample: this sample consisted of 480 schizophrenia cases and 479 controls (Zhang et al., 2014). (3) The Jewish-Israeli sample: this sample comprised 107 schizophrenia families with a total of 331

individuals (Alkelai et al., 2012). (4) The Arab-Israeli sample: this sample contained 58 schizophrenia families with a total of 198 subjects (Alkelai et al., 2011). (5) The European sample: this sample comprised 794 families with 2740 individuals (Aberg et al., 2013). (6) The African sample: this sample consisted of 438 families with 1262 individuals (Aberg et al., 2013). (7) The Asian sample: this sample contained 579 families with 2296 subjects (Aberg et al., 2013). All replication samples showed no overlap with our screening PGC1 samples. In total, 15,217 schizophrenia cases, 20,429 controls and 1976 schizophrenia families were included in the analysis. All studies were conducted under the appropriate ethical approvals, and written informed consent was obtained from all subjects.

2.2. Healthy subjects for expression quantitative loci (eQTL) analysis

To identify the potential eQTL single-nucleotide polymorphisms (SNPs) for *IRF3*, we used two well-characterized gene expression databases. The first expression database is BrainCloud (Colantuoni et al., 2011) (http://braincloud.jhmi.edu/). The BrainCloud is comprised of 261 postmortem dorsolateral prefrontal cortex (DLPFC) of normal subjects, including 113 Caucasian subjects and 148 African American individuals across the lifespan. The raw genotype data was extracted from BrainCloud, and the expression data and demographic information such as RIN, race, sex, and age were also obtained. The samples were initially divided into two groups (prenatal and postnatal), and the postnatal samples were further divided according to their ethnicities (Caucasian and African American). The statistical analysis was conducted using linear regression, with RIN, sex, and age as covariates, and race was also included when analyzing prenatal samples.

The second database is from Genevar (Yang et al., 2010) (http://www.sanger.ac.uk/resources/software/genevar/). Among Genevar (there are several datasets in Genevar), we used the European data set (CEU, Caucasians living in Utah USA, a total of 109 subjects) from Stranger et al.'s study (2012), which correlated genome-wide gene expression in lymphoblastoid cell lines with SNPs located in the region *cis* to the genes. The mRNA quantification and correlation between expression level and genotype can be found in the original study (Stranger et al., 2012).

2.3. SNP selection and genotyping

SNP selection was based on three criteria. First, tagging SNPs. The whole SNPs within *IRF3* genomic region (chr19: 50152826-50179132) were downloaded from Caucasians in 1000-Human-Genome and the Haploview program (version 4.1, Broad Institute of MIT and Harvard, Cambridge, MA) was applied to test the linkage disequilibrium (LD) between paired SNPs, to define the haplotype blocks and to select the tagging SNPs using the r^2 confidence interval (CI) algorithm (Barrett et al., 2005). Second, eQTL SNPs. By utilizing mRNA expression data from BrainCloud, we screened potential *cis* SNPs (a total of 11 SNPs in this region were available in BrainCloud) within 50 kb to *IRF3* and analyzed their associations with *IRF3* expression, and significant eQTL SNPs ($P < 0.05$) in both Caucasians and African Americans samples were chosen (Colantuoni et al., 2011). Third, potential functional SNPs. These SNPs might affect protein structure, mRNA expression and alternative splicing of the gene, such as non-synonymous SNPs and 5′-UTR SNPs. In total, eleven SNPs were selected for screening in the PGC1 sample, and top significant SNPs were further analyzed in additional samples. The LD map of the eleven SNPs in Caucasians is shown in Fig. S2 and the SNP information is shown in Table 1. For the genotyping in replication samples, we used the SNaPshot