Schizophrenia Research xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Schizophrenia Research

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

Integrated analysis supports *ATXN1* as a schizophrenia risk gene

Jiewei Liu [a,b], Bing Su [a,*]

[a] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, Yunnan, China
[b] Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming 650204, Yunnan, China

ARTICLE INFO

Article history:
Received 16 March 2017
Received in revised form 27 July 2017
Accepted 8 October 2017
Available online xxxx

Keywords:
Schizophrenia
Protein-protein interaction
ATXN1
ZNF804A
MAPK signaling pathway
Prefrontal cortex

ABSTRACT

Protein-protein interaction (PPI) is informative in identifying hidden disease risk genes that tend to interact with known risk genes usually working together in the same disease module. With the use of an integrated approach combining PPI information with pathway and expression analysis as well as genome-wide association study (GWAS), we intended to find new risk genes for schizophrenia (SCZ). We showed that *ATXN1* was the only direct PPI partner of the know SCZ risk gene *ZNF804A*, and it also had direct PPIs with other 18 known SCZ risk genes. *ATXN1* serves as one of the hub genes in the PPI network containing many known SCZ risk genes, and this network is significantly enriched for the MAPK signaling pathway. Further gene expression analysis indicated that *ATXN1* is highly expressed in prefrontal cortex, and SCZ patients had significantly decreased expression compared with healthy controls. Finally, the published GWAS data supports an association of *ATXN1* with SCZ as well as other psychiatric disorders though not reaching genome-wide significance. These convergent evidences support *ATXN1* as a promising risk gene for SCZ, and the integrated approach serves as a useful tool for dissecting the genetic basis of psychiatric disorders.

1. Introduction

Schizophrenia (SCZ) is one of the severe mental disorders with high heritability (Sullivan et al., 2003). Although many SCZ risk genes have been identified by the commonly used methods including genome-wide association studies (GWASs) (Hamshere et al., 2013; Ikeda et al., 2013; Lencz et al., 2013; Li et al., 2011; O'Donovan et al., 2008; Purcell et al., 2009; Rietschel et al., 2012; Ripke et al., 2013; Schizophrenia Psychiatric Genome-Wide Association Study Consortium, 2011; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Shi et al., 2011; Stefansson et al., 2009; Steinberg et al., 2011; Wong et al., 2014; Yue et al., 2011), linkage and association analysis of candidate genes (Allen et al., 2008; Lewis et al., 2003; Ng et al., 2009; Sun et al., 2008), exon sequencing (Fromer et al., 2014; Purcell et al., 2014) and so on, our knowledge about genetic risk factors of SCZ is still extending. Protein-protein interactions (PPIs) are binary relationships of proteins in cells. Proteins interacting with each other tend to take part in the same cellular process. According to the disease module hypothesis, proteins involved in the same disease process tend to interact with each other to form disease modules (Feldman et al., 2008; Goh et al., 2007). Such properties make PPIs highly informative for identifying hidden disease risk genes (Kohler et al., 2008; Oti et al., 2006).

ZNF804A (zinc finger protein 804A) was the first risk gene identified by GWAS (O'Donovan et al., 2008), and follow-up studies had confirmed its involvement in SCZ in many aspects (Hess and Glatt, 2014). We therefore speculate that any proteins directly interacting with *ZNF804A* may also serve as risk genes for SCZ. With the use of multiple PPI databases (HRRD, HINT and IntAct) (Keshava Prasad et al., 2009) (Das and Yu, 2012; Orchard et al., 2014), we found that the *ATXN1* gene (Ataxin-1) served as the only PPI partner of *ZNF804A*, suggesting its potential involvement in SCZ susceptibility.

ATXN1 is responsible for the neurodegenerative disease—spinocerebellar ataxia type 1 (SCA1, MIM164400). Patients have difficulties with coordination, balance, speech, swallowing, muscle and cognitive performance such as processing, learning and remembering information (https://ghr.nlm.nih.gov/condition/spinocerebellar-ataxia-type-1). A mutational mechanism in *ATXN1* with expansion of CAG repeat unit near the amino terminus has been characterized in SCA1 patients. The normal allele contains 6–39 CAG repeats while a defective version contains 40–81 repeats (Zoghbi and Orr, 1995). Several small sample studies had investigated the association between *ATXN1* and SCZ (http://www.szgene.org/geneoverview.asp?geneid=354), and the results however remained inconclusive. There were five reported case-control studies of *ATXN1* in populations of European ancestry (Culjkovic et al., 2000; Joo et al., 1999; Li et al., 1999; Morris-Rosendahl et al., 1997; Pujana et al., 1997). Two studies showed positive results while the other three were negative. In addition, four family-based association studies of *ATXN1* also showed inconsistent results (one positive and three negative) (Fallin et al., 2005; Li et al., 1999;

* Corresponding author at: Kunming Institute of Zoology, Chinese Academy of Sciences, 32 East Jiao-Chang Rd., Kunming 650223, Yunnan, China.
E-mail address: sub@mail.kiz.ac.cn (B. Su).

https://doi.org/10.1016/j.schres.2017.10.010
0920-9964/

Lin et al., 2009; Pujana et al., 1997). The inconsistency of the reported data was possibly due to the limited sample sizes of these studies (the numbers of cases were less than 350).

The direct PPI network of a set of genes may imply their involvement in a specific disease process. We thus hypothesize that *ATXN1* may participate in SCZ etiology through direct interaction(s) with known SCZ risk genes such as *ZNF804A*. We utilized an integrated approach by combining analyses of PPI network, molecular pathway, gene expression and genetic association, we showed that *ATXN1* directly interacts with many known SCZ risk genes including *ZNF804A*, and it significantly participated in the SCZ disease network.

2. Materials and methods

2.1. Protein-protein interaction analysis

We firstly extracted PPI pairs of *ZNF804A* from the HPRD database (Keshava Prasad et al., 2009). For validation, we also extracted PPI information of *ZNF804A* from HINT (Das and Yu, 2012) and IntAct (Orchard et al., 2014). It turned out that *ATXN1* is the only PPI partner of *ZNF804A* and this result is consistent among all three databases. We next constructed PPI information of *ATXN1* using the HPRD database. We excluded the PPI pairs with missing protein information. The PPI network was visualized in the Cytoscape platform (Table S1 and Fig. 1) (Shannon et al., 2003). Genes were annotated with SCZ and non-SCZ categories. The SCZ gene list was download from the SZDB database (Wu et al., 2016) (http://www.szdb.org/). Furthermore, the SCZ genes were classified into six groups based on the approaches used in identifying these genes including GWASs (Hamshere et al., 2013; Ikeda et al., 2013; Lencz et al., 2013; Li et al., 2011; O'Donovan et al., 2008; Purcell et al., 2009; Rietschel et al., 2012; Ripke et al., 2013; Schizophrenia Psychiatric Genome-Wide Association Study Consortium, 2011; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Shi et al., 2011; Stefansson et al., 2009; Steinberg et al., 2011; Wong et al., 2014; Yue et al., 2011), exon sequencing (Fromer et al., 2014; Purcell et al., 2014), Sherlock integrative analysis (Luo et al., 2015), convergent functional genomics (Ayalew et al., 2012), linkage and association analysis (Allen et al., 2008; Lewis et al., 2003; Ng et al., 2009; Sun et al., 2008) and genes affected by copy number variations (Luo et al., 2014a). We labeled the SCZ genes as "multiple evidences" if they were reported by more than two different approaches (Fig. 1 and Table S2).

We carried out a simulation to investigate if *ATXN1* has unusually high direct PPIs with SCZ genes. We took the SCZ genes as the background genes, and we extracted direct interactions among them. Degree parameter of each SCZ gene was calculated by the Cytoscape plug-in Network Analyzer (Doncheva et al., 2012). The degree distribution was plotted using R.

In order to explore *ATXN1*'s relationship with network formed by SCZ susceptibility genes. We selected two sets of reported SCZ genes used in a previous study (Luo et al., 2014c). The first dataset contained 34 genes that reached genome-wide significance in recent SCZ GWASs (Hamshere et al., 2013; Ikeda et al., 2013; Li et al., 2011; Luo et al., 2014b; O'Donovan et al., 2008; Purcell et al., 2009; Rietschel et al., 2012; Schizophrenia Psychiatric Genome-Wide Association Study Consortium, 2011; Shi et al., 2011; Stefansson et al., 2009; Steinberg et al., 2011; Yue et al., 2011) (Table S3). The second covered 42 genes identified by Convergent Functional Genomics (Ayalew et al., 2012) (Table S3). TCF4 is the only gene shared between the two gene sets. We used DAPPLE (Disease Association Protein–Protein Link Evaluator, http://www.broadinstitute.org/mpg/dapple/dapple.php) to explore whether *ATXN1* participates in the network formed by the SCZ genes (Rossin et al., 2011). DAPPLE uses the InWeb PPI data to construct PPI network, and InWeb contains 169,801 high-confidence PPI items prioritized from HINT, BIND, IntAct and PPrel protein interaction databases. Permutation test (1000 times) was performed to evaluate the significance of the PPI network. This analysis was implemented in the GenePattern platform (http://software.broadinstitute.org/cancer/software/genepattern) that contains integrated data analysis toolkits, and DAPPLE is one of them.

2.2. Pathway enrichment analysis of ATXN1 PPI network

We used an online tool (WebGestalt—WEB-based gene set analysis toolkit, http://www.webgestalt.org/) (Wang et al., 2013). The genes of the *ATXN1* PPI network were used as queries for WebGestalt (Table S1). We used two databases (KEGG and Wikipathway) to perform pathway enrichment analysis. The significance of pathway enrichment analysis was set to $p < 0.01$ (after adjustment for multiple tests).

2.3. Expression analysis

The tissue enrichment expression analysis of *ATXN1* was retrieved from an online resource, BioGPS (http://biogps.org/) (Wu et al., 2009).

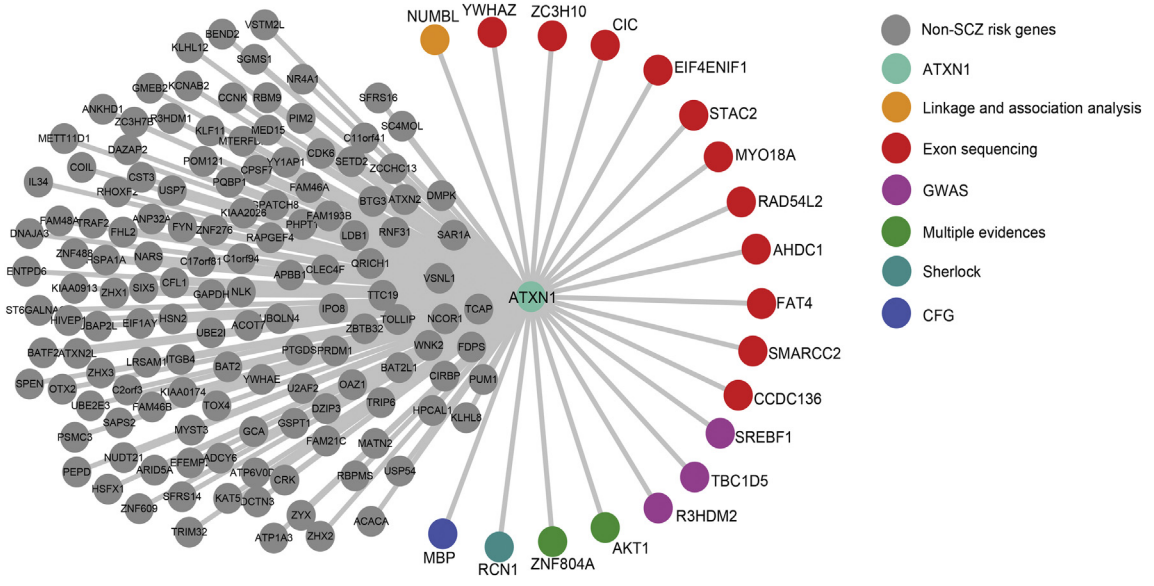

Fig. 1. The PPI network of *ATXN1*. The known SCZ genes were color-labeled according to the approaches used in identifying them (details are provided in the method).

Please cite this article as: Liu, J., Su, B., Integrated analysis supports *ATXN1* as a schizophrenia risk gene, Schizophr. Res. (2017), https://doi.org/10.1016/j.schres.2017.10.010