

SCIENTIFIC REPORTS



OPEN

Altered expression of mRNA profiles in blood of early-onset schizophrenia

Received: 11 March 2015
Accepted: 20 October 2015
Published: 06 January 2016

Yong Xu^{1,*}, Yin Yao Shugart^{2,*}, Guoqiang Wang³, Zaohuo Cheng³, Chunhui Jin³, Kai Zhang³, Jun Wang³, Hao Yu^{4,5}, Weihua Yue^{4,5}, Fuquan Zhang³ & Dai Zhang^{4,5,6}

To identify gene expression abnormalities in schizophrenia (SZ), we generated whole-genome gene expression profiles using microarrays on peripheral blood mononuclear cells (PBMCs) from 18 early-onset SZ cases and 12 controls. We detected 84 transcripts differentially expressed by diagnostic status, with 82 genes being upregulated and 2 downregulated. We identified two SZ associated gene coexpression modules (green and red), including 446 genes. The green module is positively correlated with SZ, encompassing predominantly up-regulated genes in SZ; while the red module was negatively correlated with disease status, involving mostly nominally down-regulated genes in SZ. The olfactory transduction pathway was the most enriched pathways for the genes within the two modules. The expression levels of several hub genes, including AKT1, BRCA1, CCDC134, UBD, and ZIC2 were validated using real-time quantitative PCR. Our findings indicate that mRNA coexpression abnormalities may serve as a promising mechanism underlying the development of SZ.

Schizophrenia (SZ) is a severe chronic mental disorder affecting about 1% of the population worldwide. Generally arising in late adolescence, it profoundly disrupts a few key traits of human cognition and personality including language, thought, perception, emotional affect, and sense of self. SZ patients tend to first present with overt symptoms during late adolescence or early adulthood and it has been postulated that this developmental stage represents a “window of vulnerability”. When the disease manifests before age 18, it is defined as early-onset SZ (EOS), a subcategory of SZ associated with more familial vulnerability and poor outcomes¹.

Despite SZ’s wide prevalence and debilitating nature, little is known about its pathogenesis. Clinicians therefore tend to rely on clinical symptoms for diagnosis and for evaluating the progress and treatment response throughout the course of the disease. It has been hypothesized that the gene expression is the most fundamental level at which the genotypes critically influence the SZ phenotypes.

It is almost impractical to obtain biopsied brain tissue from SZ patients for the development of a molecular signature that may assist a diagnosis. In this regard, peripheral blood mononuclear cells (PBMCs) can be easily collected from patients and followed longitudinally with gene expression analyses, which may provide a way of identifying the signatures of clinical subtypes, their prognosis and treatment response.

¹Department of Psychiatry, First Hospital /First Clinical Medical College of Shanxi Medical University, Taiyuan, China. ²Unit on Statistical Genomics, Division of Intramural Research Program, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland, United States of America. ³Wuxi Mental Health Center, Nanjing Medical University, Wuxi, Jiangsu Province, China. ⁴Institute of Mental Health, Sixth Hospital, Peking University; Beijing 100191, China. ⁵Key Laboratory of Mental Health, Ministry of Health & National Clinical Research Center for Mental Disorders (Peking University), Beijing, 100191, China. ⁶Peking-Tsinghua Center for Life Sciences/PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, 100871, China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to W.Y. (email: dryue@bjmu.edu.cn) or F.Z. (email: zhangfq@njmu.edu.cn)

To detect biomarkers for SZ and other psychiatric disorders, a few earlier studies profiled gene expression in peripheral blood^{2–7}. However, the results varied across studies. To date, a consistent pattern of alterations has not been established^{3,8}.

The expressions of transcripts are influenced by many environmental factors, including medication. Most previous studies were conducted using patients under drug treatment or with a history of pharmacotherapy, which makes it impossible to preclude the potential effect of antipsychotic therapy. However, several studies have been conducted on PBMCs in drug-naïve participants. Craddock *et al.*⁹ used blood T cell derived RNA and compared the gene expression from six minimally treated or first-episode SZ patients (age 31.6 ± 14.1) with controls and identified 399 differentially expressed (DE) probes. Of these, 320 (80%) probes were decreased in SZ and 79 were increased. Takahashi *et al.*¹⁰ identified 792 DE probes, with 256 probes being downregulated and 536 probes (68%) being up-regulated in blood of antipsychotics-free SZ patients (age 31.8 ± 11.4) compared with controls. Recently, Kumarasinghe *et al.*¹¹ detected 416 (67%) down-regulated and 208 up-regulated genes in blood of treatment-naïve patients (age 36.1 ± 14.8) when compared to controls. However, these studies showed that the patients had diverse disease durations and their age spanned a wide range (> 20 years), which will certainly affect the results. Furthermore, it remains controversial concerning the global direction of changes in genome expression for SZ.

It is well known that genes tend to work together to perform its function, and functionally related genes tend to be coexpressed^{12,13}. It is tempting to apply the gene coexpression network analysis to results interpretation. Gene coexpression networks encapsulate the activity of multiple regulatory systems and hold the potential to highlight specific molecular mechanisms for disease¹⁴. Differential coexpression refers to variations in gene–gene correlation between two sets of phenotypically distinct samples¹⁵. Gene–gene correlation may change without affecting differential expression, indicating that a gene may alter its regulatory pattern that would be missed by traditional differential expression analyses. Differential coexpression analysis (DCEA), which aims to find gene modules with different connectivity (correlations) in the disease state, offers a more powerful approach for elucidating transcriptome patterns and dysfunction of gene expression underlying phenotypic changes. Several studies conducted gene coexpression analysis for mental health disorders^{16–20}. One commonly used method is weighted gene coexpression network analysis (WGCNA)²¹.

In this work, we aimed to identify potential biological markers for SZ using blood-based gene expression profiles from a cohort of EOS and healthy controls. We detected a panel of individually misexpressed genes implicated in SZ, suggesting an upregulation trend of gene expression in SZ. Further, we identified two clusters of coexpressed genes associated with SZ. Further analysis revealed intensive interactions among the genes within the two clusters and that several hub genes may represent new causal candidate genes for SZ.

Results

mRNA profiles in SZ cases compared with controls. In this study, we used the microarray platform and detected a total of 8594 mRNAs. Cluster analysis is shown in Supplementary Figure S1. DE mRNAs were identified through FC and P value filtering. With P values adjusted using Bonferroni correction ($FC \geq 2$ and $P_{\text{adjusted}} < 0.05$), 84 mRNAs were DE, of which, 82 were upregulated and 2 downregulated in SZ patients compared with healthy control subjects (Fig. 1, Supplementary Table S1). Major GO categories of the 84 DE genes are shown in Supplementary Figure S2. Gene ontology analysis identified 31 categories with P value < 0.001, three of which reached FDR significance (plasma membrane, nucleosome, and nuclear nucleosome) (Fig. 2A, Supplementary Table S2). KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis identified 30 non-directional enriched pathways (Fig. 2B, Supplementary Table S3).

Assuming a sample size in each group of 15 and a FDR of 0.05, our dataset offers a power of 83.3% to detect DE genes (Supplementary Figure S3). Since our samples comprised of 18 cases and 12 controls, the power of our dataset may be slightly lower the estimated power.

Coexpression analysis. Genes with coefficient of variation of expression levels less than 0.1 were filtered out. A total of 2484 out of 8518 genes were retained in the coexpression analysis. The cluster dendrogram is shown in Fig. 3. There were 7 coexpressed modules, among which 2 modules (green and red) were significantly associated with SZ (Table 1), including a total of 446 genes (Supplementary Table S4). The green module is positively correlated with SZ, meaning that the genes in this module are predominantly up-regulated in SZ cases. In contrast, the red module was negatively correlated with disease status, involving mostly down-regulated genes in SZ. Topological Overlap Matrix (TOM) plot of the network connections is shown in Supplementary Figure S4, and gene significance across modules is shown in Supplementary Figure S5.

Notably, 55 out of 84 DE genes were among the 238 green module genes. All of the 55 genes were up-regulated in SZ compared with controls. Given that we observed a total of 2484 genes in the coexpression modules, we would like to indicate the strength of the statistical rigor for our finding ($P < 2.2E-16$).

Inner module structure of the green and red modules is shown in Fig. 4. Genes with high intra-modular connectivity are informally referred to as intra-modular hub genes. Our results indicated that CCDC134

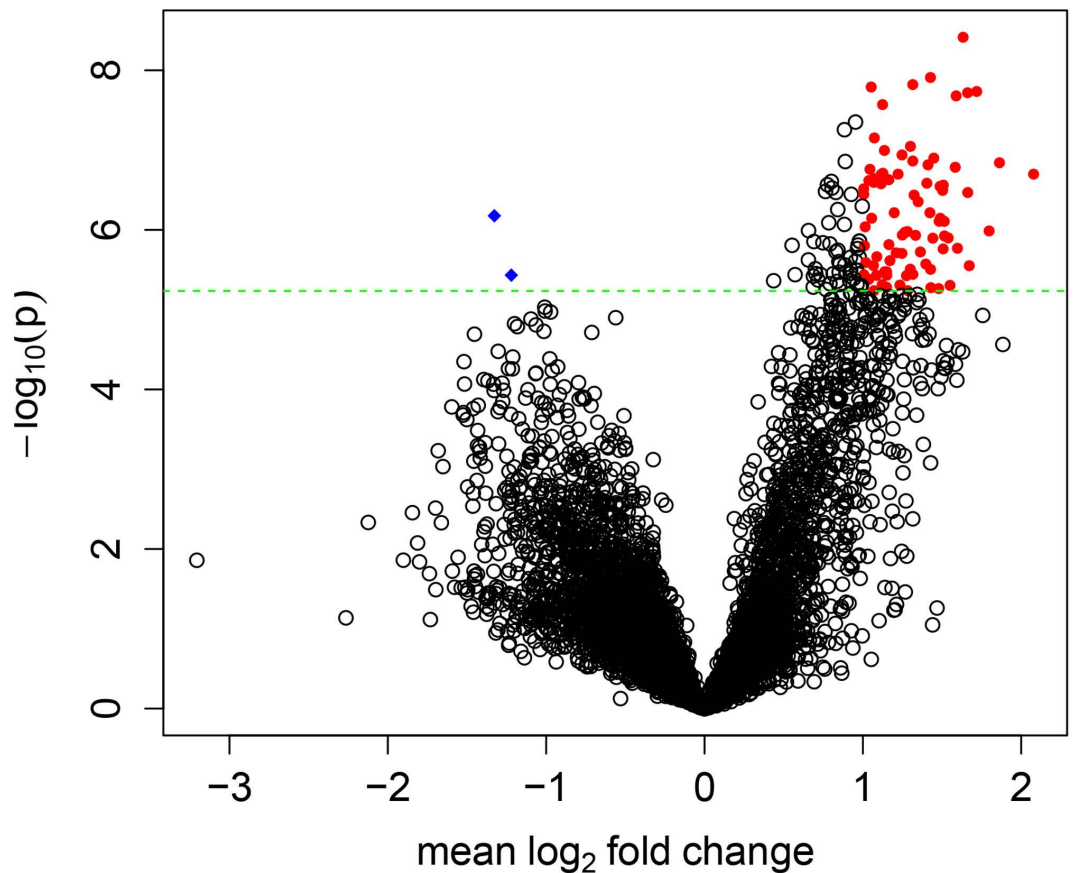


Figure 1. Volcano plot of mRNA expressions. Plotted along the x-axis is the mean of \log_2 fold-change, along the y-axis the negative logarithm of the p-values. Red denotes the 82 up-regulated genes, blue the 2 down-regulated genes. The horizontal green line is the negative logarithm the Bonferroni-adjusted P-value threshold ($5.82E-6$).

and UBD are hub genes in the green module. While within the red module, C10orf67, TNIP3, and ZIC2 are hub genes.

Pathway analysis of the green and red module. Protein–protein interaction (PPI) analysis shows the existence of complex interactions among the 446 genes within the green and red modules (Fig. 5A). AKT1, CHEK1 and BRCA1 were hub genes in the network. GO analysis shows that 10 biological processes are enriched in the 446 module genes (Supplementary Table S5). KEGG-based pathway analysis indicates that olfactory transduction and protein digestion and absorption are two enriched pathways (Fig. 5B).

Real-time quantitative PCR (RT-qPCR) analysis of hub genes. We evaluated blood expression levels of AKT1, BRCA1, CCDC134, UBD, and ZIC2 in 48 healthy controls, and 30 SZ patients before and after a 12-week following-up treatment using RT-qPCR. Our data show that GAPDH was expressed at a stable level across the samples. At the threshold of $FDR < 0.05$, AKT1 and BRCA1 were up-regulated in SZ cases compared with controls (Fig. 6, Supplementary Table S6); while CCDC134, UBD, and ZIC2 were down-regulated after the 12-week treatment (Fig. 6, Supplementary Table 7).

Discussion

The transcriptome sits between environmental influence and the genetic susceptibility to SZ and thus may serve as a bridge between certain endophenotypes and the genetic changes that lead to the disorder. To gain a better understanding of the neuropathology of SZ, therefore, we sought to conduct mRNA profile examinations in a cohort of drug-free teenaged patients with first-onset SZ. The study group was homogeneous in age, all were students, and had no other major physical diagnosis. Most importantly, all the patients were untreated, precluding any confounding effects of therapeutic drugs. Because of the fact that the patients were in the early stage of disease and given their ages, their brains were still developing. We therefore speculate that the alterations in transcription identified in these patient cohorts are likely to be relevant to the disease process.

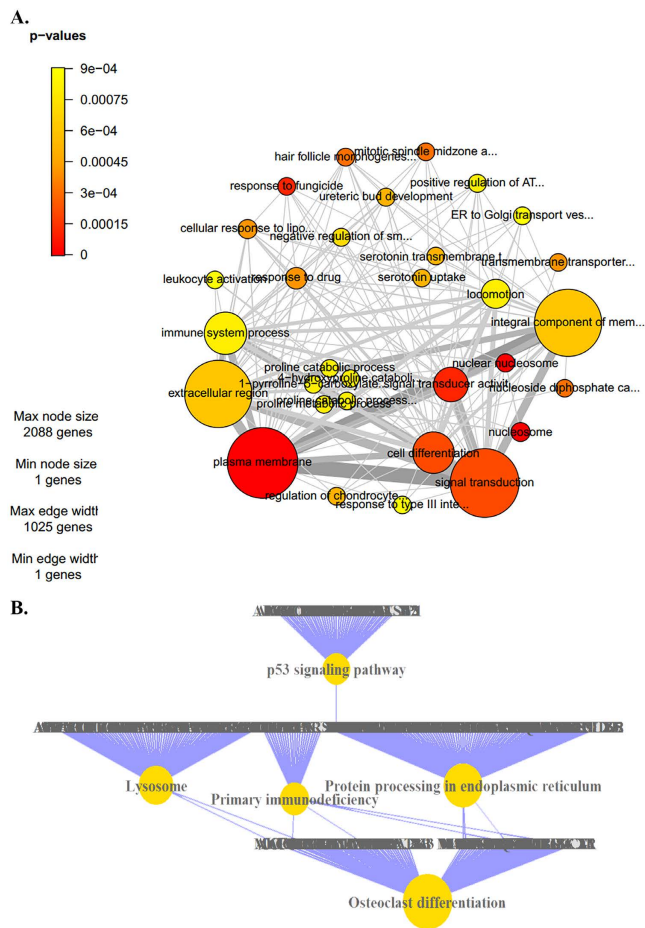


Figure 2. Pathway analysis of the genes. Gene sets that are significant ($p < 0.001$) in the non-directional class from the GO enrichment analysis. Only gene sets with 5 or more genes were retained. The network plot shows the relation between gene sets by connecting those that share member genes. The color of nodes indicates P values. The thickness of the edges corresponds to the number of shared genes. The size of the nodes corresponds to the size of the gene sets.

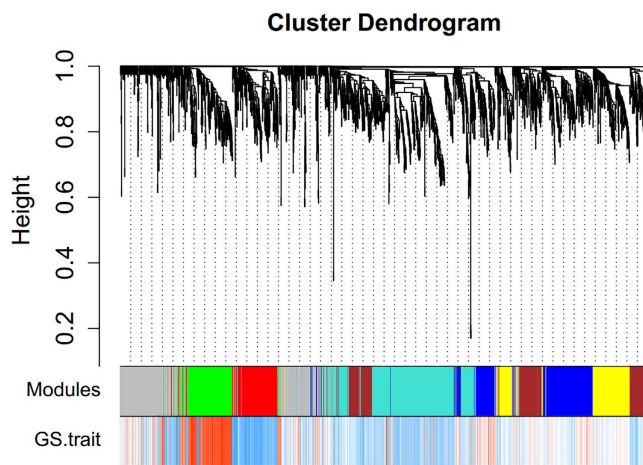


Figure 3. Hierarchical cluster tree of the 2484 genes. The color bands provide a simple visual comparison of module assignments (branch cuttings) based on the dynamic hybrid branch cutting method. The first band shows the results from the automatic single block analysis and the second color band visualizes the gene significance measure: “red” indicates a negative correlation with disease, and “blue” indicates a positive correlation.

Module	Genes	Coefficient	P
green	238	0.766	8.28E-7
red	208	-0.579	8.01E-4
turquoise	578	-0.345	0.062
brown	316	-0.200	0.290
blue	372	0.062	0.747
yellow	259	-0.037	0.845
grey	513	0.009	0.962

Table 1. Association of module eigengene with disease.

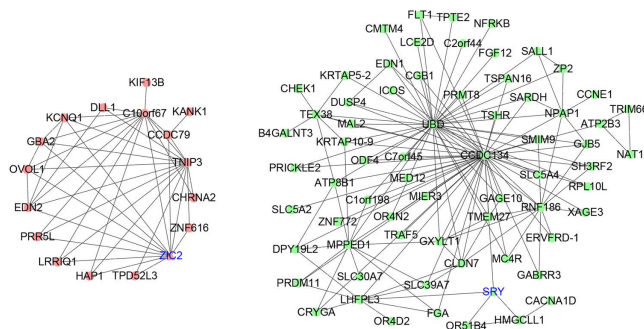


Figure 4. Intra-modular connectivity of the green and red modules. Green nodes denote genes in the green module; red nodes denote red module genes. Triangle nodes with blue labels are transcriptional factors.

The period of late adolescent development is of particular interest to psychiatry research, as this time window corresponds to the age of onset of major neuropsychiatric disorders, especially SZ. We focused specifically on this period, with the goal of identifying genes whose expression is altered during this period. We detected predominant up-regulation of mRNAs in the SZ group, with 82 out of 84 DE genes being increased. This finding was consistent with those of Takahashi *et al.*¹⁰ regarding the change direction. This observation seems valuable considering that transcript studies are often plagued with uncertainties, inconsistencies and conflicts due to fluctuations of gene expression and a myriad influencing factors³.

Gene ontology analysis of the 84 DE genes showed that these genes were enriched in signal transduction, multicellular organismal development, metabolic process, ion transport, cell cycle, and cell differentiation. Among the 84 DE genes, solute carrier family 18 (vesicular monoamine transporter), member 1 (SLC18A1) and cytotoxic T-lymphocyte-associated protein 4 (CTLA4) have been previously implicated in SZ.

Vesicular monoamine transporters (VMATs) are involved in the packaging of dopamine, serotonin, adrenalin, and noradrenalin from the cytoplasm to their storage vesicles in presynaptic terminals. Dysregulation of monoamines has been long postulated to play an important role in the etiology of SZ and other neuropsychiatric disorders. SLC18A1 is one of the two isoforms of VMATs, playing a critical role in the maintenance of monoaminergic endocrine systems. SLC18A1 is mapped to chromosome 8p21.3, a locus with strong evidence for linkage with SZ^{22,23}. SLC18A1 has been reported to be associated with SZ²⁴⁻²⁷. In this study, we detected an increased level of expression of SLC18A1 in SZ patients, providing further support for its possible involvement into SZ.

CTLA4 is involved in establishing and maintaining peripheral T-cell tolerance, which controls T-cell activation and reactivity. The CTLA4 gene is located on chromosome 2q33, and has been suggested as a candidate gene for conferring susceptibility to autoimmune disease²⁸. Numerous studies have demonstrated that SZ patients showed a T-cell-associated dysfunction²⁹. Several studies have reported the association of CTLA4 variants with SZ³⁰⁻³³. Furthermore, significantly increased expression of CTLA4 was observed³², consistent with the results from our study.

DCEA identified two modules relevant to SZ. The green module is most striking, in that it has a high correlation with the disease, and involves most up-regulated genes (67.1%). This observation further suggested that most DE genes may be functionally related. The red module is an additional finding coming from the coexpression analysis. Although most genes within the red module are down-regulated, none of them reached the genome-wide significance level. These genes were missed by traditional differential expression analysis. However, these genes may also contribute to the development of the disease. The

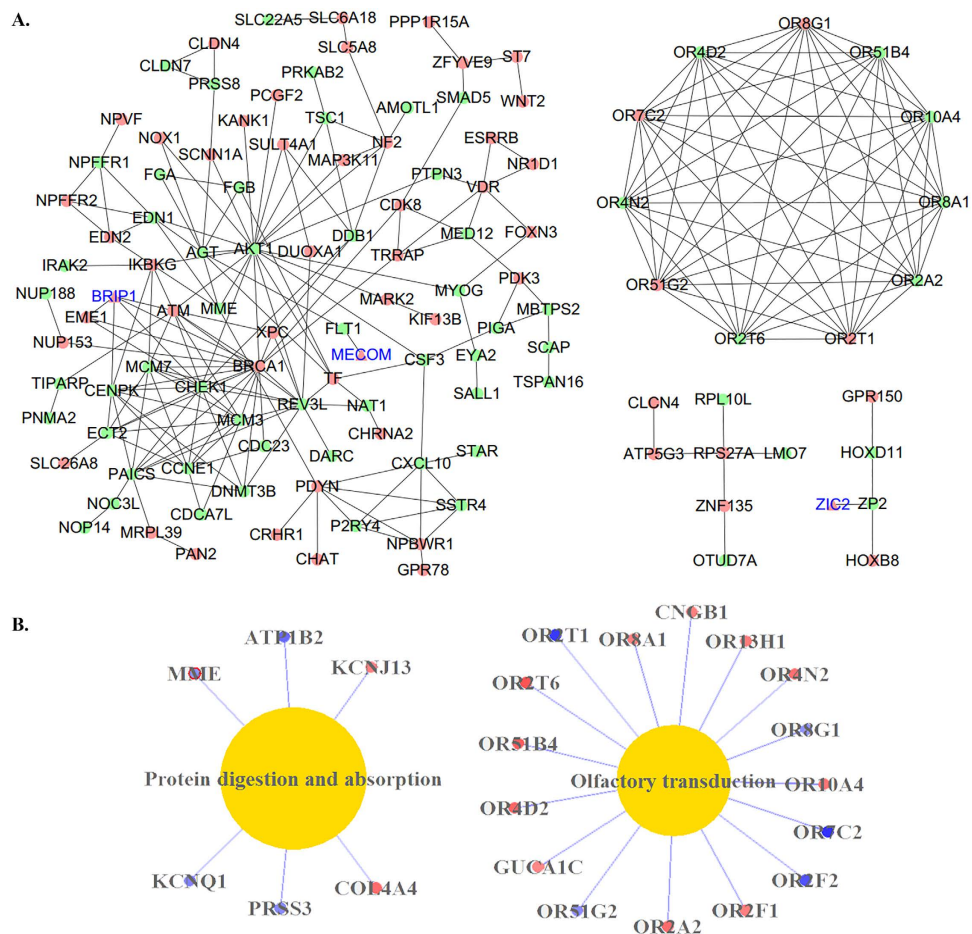


Figure 5. PPI and pathway analysis of the genes within the green and red modules. (A) PPIs among the module genes (networks with nodes >4 were included). Green nodes denote genes in the green module; red nodes denote red module genes. Triangle nodes with blue labels are transcriptional factors. (B) Enriched KEGG pathways of the module genes. Red nodes denote up-regulated genes, while blue nodes denote down-regulated genes.

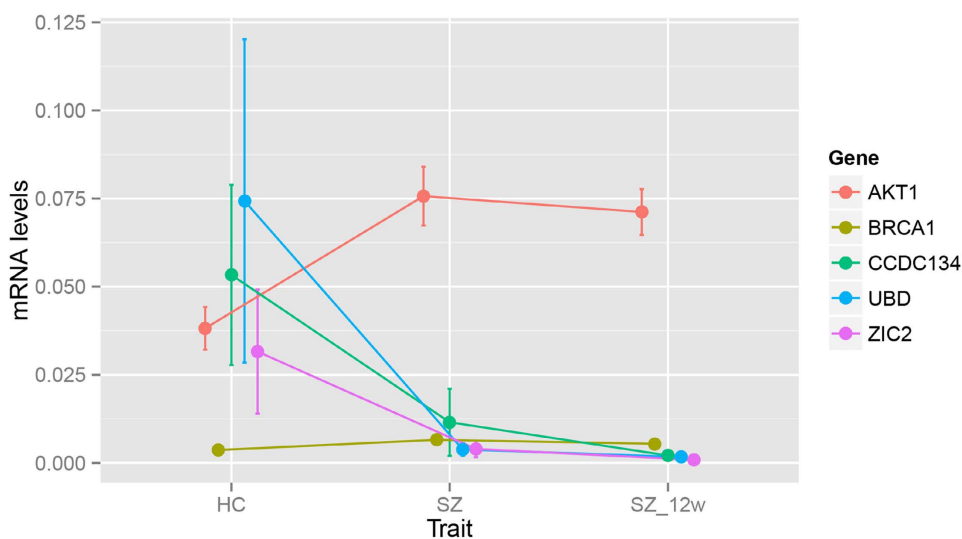


Figure 6. Expression levels (log2 transformed) of several key genes via RT-qPCR analysis. HC: healthy control; SZ: untreated schizophrenia patients; SZ_12w: schizophrenia patients after 12-week treatment. Vertical bars denote 95% confidence interval of expression levels for each gene.

green and the red modules involve two distinctive clusters of genes with opposite expression trends and therefore association directions. Nonetheless, the intensive PPIs among the genes within two modules indicate that the two modules are interconnected and act jointly.

Coexpression modules highlight several hub genes, including *CCDC134*, *UBD*, *C10orf67*, *TNIP3*, and *ZIC2*, while PPI analysis indicates *AKT1*, *CHEK1* and *BRCA1* as hub genes. All of these hub genes are expressed in brain (Supplementary Figure S6). Since the majority of other genes are interacted with these hub genes, we postulate that most genes in the PPI network should be expressed in brain too. Among the hub genes, *ZIC2*, *AKT1*, *CHEK1*, and *BRCA1* have been implicated in SZ, while *CCDC134*, *UBD*, *C10orf67*, and *TNIP3* may be promising biomarkers warrant further investigation.

ZIC2 is a member of the *ZIC* family of zinc finger proteins, and functions as a transcriptional repressor with critical roles in neural development³⁴. It has been suggested that the mice heterozygous for the hypomorphic mutation in *Zic2* mimic the SZ-like phenotypes and *ZIC2* mutations were associated with SZ^{35,36}. *ZIC2* inhibits Wnt/ β -catenin protein signaling pathway that is relevant to the pathogenesis of SZ^{37,38}. However, no significant difference between SZ cases and controls was observed in both our microarray dataset ($FC = 0.43$, $P_{\text{Bonferroni}} = 1.00$) and the RT-PCR analysis ($FC = 0.13$, $FDR = 0.074$).

A common function among *CHEK1*, *BRCA1*, and *AKT1* is that they are relevant to cell cycle. *CHEK1* and *BRCA1* have been reported to be downregulated in SZ³⁹. The impairment of *AKT/GSK3* signaling pathway has been implicated in SZ and decreased *AKT1* protein levels and phosphorylation have been observed in lymphocytes and brains of individuals with SZ^{40,41}. Several studies also detected the association of *AKT1* genetic variants with SZ (reviewed in Ref. 42).

The expression levels of *AKT1*, *BRCA1*, *CCDC134*, *UBD*, and *ZIC2* were validated in the RT-PCR analysis, with *BRCA1* and *AKT1* being upregulated in the patients; while after the 12-week treatment, the other three genes, *CCDC134*, *UBD*, and *ZIC2* were slightly down-regulated. This result indicated that these genes may be trait-relevant rather than state-relevant.

Although there was a paucity of overlaps of the 84 DE genes with well-known candidate genes for SZ, several enriched pathways in our mRNA dataset have been implicated in SZ, including the p53 signaling pathway^{43,44}, the chemokine signaling pathway, and the MAPK signaling pathway⁴⁵. Interestingly, a new pathway, the olfactory transduction pathway, was identified in the 446 module genes in our study. Impairments of olfactory system in schizophrenia patients have been well-documented^{46,47}. It has been suggested that dysfunction in the central and peripheral olfactory system in SZ is ubiquitous, which may be a useful marker of SZ⁴⁷. Our study seemed to support the alteration of olfactory pathways in the blood of SZ patients.

Several limitations in this study need to be noted. Given the rarity of drug-naïve patients with secure diagnoses of EOS, the sample in the expression experiments was relatively small. A caveat in mind was that the mRNA results were derived from blood sample, thus, due care should be taken when extrapolating these results into the brain. It has been suggested that peripheral biomarkers may in part mirror similar pathological processes in the brain, in part represent a distinct molecular changes in blood that is highly specific to the primary pathophysiology, or reflect the molecular response of the peripheral cells secondary to the disease⁴⁸. Thus, abnormalities in blood, whether or not they are mimicked in brain, are potential indicators of disease pathology.

Our study identified two clusters of SZ associated coexpressed modules, which may help unravel mechanisms underlying the disease. Several genes within the modules may be novel promising biomarkers for SZ.

Methods

Subjects. All participants were unrelated Han Chinese recruited from the north of China. Consensus diagnoses were made by at least two experienced psychiatrists independently according to the Diagnosis and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV) criteria for SZ. Patients with unanimous diagnosis were enrolled into the study.

In the microarray analysis, samples included 18 first-onset SZ patients (8 males and 10 females, aged 14.78 ± 1.70 years ranging from 10–18 years), and 12 healthy controls (6 males and 6 females, aged 14.75 ± 2.14 years ranging from 10–17 years). Healthy teenager controls were recruited from several middle schools with an interview conducted by psychiatrists; individuals with a history of mental health or neurological diseases were all excluded.

For the real-time quantitative PCR (RT-qPCR) analysis, we enrolled 38 SZ patients who were drug-free for at least one month before the enrollment; among them 30 patients (14 males and 16 females, aged 34.5 ± 11.0 years) were successfully followed-up with a 12-week period of antipsychotic treatment. The clinical effects were assessed by trained and experienced psychiatrists with the Positive and Negative Syndrome Scale (PANSS) respectively before and after 12-week treatment. All patients got clinical improvement according to the PANSS reductive ratio more than 25%. All patients participating in the study were treated with one of the oral second generation or atypical antipsychotics (SGA) and tracked for 12-week continuous medication after baseline assessments. A total of 48 healthy controls (17 males and 31 females, aged 31.6 ± 6.88 years) were recruited from local communities or were undergoing routine health check-ups. Subjects with relevant physical diseases or a history of major psychiatric disorders or suicidal behavior were excluded, and those who had a first-degree relative with a history of severe mental disorder or suicidal behavior were also excluded.

There was no significant difference in gender or age between SZ cases and controls in the two cohorts of samples (Supplementary Table S8). Total RNA was isolated from peripheral blood mononuclear cells (PBMCs) using TRIzol (Invitrogen; USA) with on-column DNase I treatment as described by the manufacturer. Blood samples of each participant were collected on early morning before breakfast. Following the diagnosis, blood samples of the schizophrenia patients were collected next morning.

The study was approved by Medical Research Ethics Committee of Shanxi Medical University, and all experiments were performed in accordance with the approved guidelines and regulations. Informed consent was signed by both the teenage participants and their parents or care takers.

mRNA profiling. Total RNA was isolated from peripheral blood mononuclear cells (PBMCs) using TRIzol (Invitrogen; USA) with on-column DNase I treatment as described by the manufacturer. The concentration and purity of isolated RNA can be measured with the NanoDrop nd-1000 UV-Vis spectrophotometer. The concentration of isolated RNA was more than 1000ng/ul, the purity of RNA were both OD260/280 and OD260/230 very closed to 2.0. The intensity of these rRNA bands on denaturing agarose gels have been used to calculate a ratio that served as an indication of RNA integrity. A 28S/18S ratio of two indicates a good quality of RNA.

Agilent Array platform was employed to perform the microarray analysis. The sample preparation and microarray hybridization were performed based upon the manufacturer's standard protocols with minor modifications. Briefly, mRNA was purified from total RNA after removal of rRNA (mRNA-ONLY™ Eukaryotic mRNA Isolation Kit, Epicentre). Then, each sample was amplified and transcribed into fluorescent cRNA along the entire length of the transcripts without 3' bias utilizing a random priming method. The labeled cRNAs were hybridized onto the Human LncRNA Array v2.0 (8 × 60 K, Arraystar). After having washed the slides, the arrays were scanned by the Agilent Scanner G2505B. Agilent Feature Extraction software (v10.7.3.1) was used to analyze acquired array images. Raw signal intensities were normalized in quantile method using GeneSpring GX v11.5.1 (Agilent Technologies), and low intensity mRNAs were filtered (mRNAs that at least 20 out of 30 samples have flags in Present or Marginal were chosen for further analysis). After normalization, the distributions of log₂-ratios among all samples were nearly the same (Supplementary Figure S7). The raw microarray data is deposited at gene expression omnibus (GEO) and are accessible through GEO Series accession number GSE54913.

A total of 17219 valid probes were detected. Data on the probe level was collapsed to genes for separate datasets prior to analysis. When two or more probes were available (this is the case for most genes), expression levels of the probes were averaged to represent the expression level of the gene. This resulted in 8594 genes. Mis-mapped genes between gene symbols and Entrez IDs were filtered, resulting in 8518 genes.

Analysis of gene expression by RT-qPCR. cDNA was synthesized using High Capacity RNA-to-cDNA Kit (Invitrogen; USA) as described by the manufacturer. The primers were listed in Supplementary Table S9. PCR was performed using a ViiA 7 Real-time PCR System (Applied Biosystems) for 10 min at 95 °C, and then 40 cycles consisting of 10 s at 95 °C, 60 s at 60 °C, 15 s at 95 °C, followed by a subsequent standard dissociation protocol to ensure that each amplicon was a single product. All quantifications were normalized to GAPDH.

Statistical analysis and in silico prediction. R⁴⁹ was used to perform the data processing and analyses of mRNA data. Student's t-test was used to compare expression levels between the patient and control groups. Power of the gene expression data was estimated using R package *ssize.fdr*^{50,51}. Clustering heat maps were plotted with *gplots*⁵². Gene expression data were filtered using R package *genefilter*⁵³. Gene ontology (GO) analysis was performed using R package *piano*⁵⁴. KEGG pathway analysis was performed using R package *GeneAnswers*^{55,56}, and GO pie chart was plotted using R packages *geneListPie*⁵⁷. Protein-protein interaction (PPI) network was retrieved via STRING v9.1⁵⁸. The networks were plotted using Cytoscape⁵⁹.

Coexpression analysis was conducted using R package WGCNA v1.43^{21,60,61}. A weighted adjacency matrix containing pair-wise connection strengths was constructed by using the soft-thresholding approach ($\beta = 30$, Supplementary Figure S8) on the matrix of pairwise correlation coefficients. A connectivity measure (*k*) per gene was calculated by summing the connection strengths with other genes. Modules are defined as clusters of densely interconnected genes; by default, they are indicated by branches of a hierarchical clustering tree using a dissimilarity measure. Then, each module is subsequently assigned a color. The gene expression profiles of each module were summarized by the module eigengene (defined as the first principal component of the module expression levels). A measure of gene significance (GS) is computed to evaluate a gene's correlation with a phenotype. The intra-modular connectivity (*k*-within) was calculated for each gene by summing the connection strengths with other module genes and dividing this number by the maximum intra-modular connectivity. The adjacency threshold for including edges in the output was set at 0.2. Eigengene-based connectivity of a gene in a module is defined as the correlation of the gene with the corresponding module eigengene.

For RT-qPCR analysis, the comparative Ct ($2^{-\Delta\Delta CT}$) method was used for the quantification of transcripts. The differences of mRNA levels between cases and controls were analyzed using the Wilcoxon

rank-sum test, and the differences of mRNA levels between schizophrenia cases before and after the treatment were analyzed using the paired Wilcoxon rank-sum test. The significance threshold was set at false discovery rate (FDR) < 0.05.

References

- Clemmensen, L., Vernal, D. L. & Steinhausen, H. C. A systematic review of the long-term outcome of early onset schizophrenia. *BMC Psychiatry* **12**, 150 (2012).
- Schwarz, E. *et al.* Validation of a blood-based laboratory test to aid in the confirmation of a diagnosis of schizophrenia. *Biomark Insights* **5**, 39–47 (2010).
- Kumarasinghe, N., Tooney, P. A. & Schall, U. Finding the needle in the haystack: a review of microarray gene expression research into schizophrenia. *Aust. N. Z. J. Psychiatry* **46**, 598–610 (2012).
- Sanders, A. R. *et al.* Transcriptome study of differential expression in schizophrenia. *Hum. Mol. Genet.* **22**, 5001–5014 (2013).
- Sainz, J. *et al.* Inflammatory and immune response genes have significantly altered expression in schizophrenia. *Mol. Psychiatry* **18**, 1056–1057 (2013).
- Gardiner, E. J. *et al.* Gene expression analysis reveals schizophrenia-associated dysregulation of immune pathways in peripheral blood mononuclear cells. *J. Psychiatr. Res.* **47**, 425–437 (2013).
- Schwarz, E. *et al.* Identification of a biological signature for schizophrenia in serum. *Mol. Psychiatry* **17**, 494–502 (2012).
- Horvath, S., Janka, Z. & Mirnics, K. Analyzing schizophrenia by DNA microarrays. *Biol. Psychiatry* **69**, 157–162 (2011).
- Craddock, R. M. *et al.* Altered T-cell function in schizophrenia: a cellular model to investigate molecular disease mechanisms. *PLoS One* **2**, e692 (2007).
- Takahashi, M. *et al.* Diagnostic classification of schizophrenia by neural network analysis of blood-based gene expression signatures. *Schizophr. Res.* **119**, 210–218 (2010).
- Kumarasinghe, N. *et al.* Gene expression profiling in treatment-naïve schizophrenia patients identifies abnormalities in biological pathways involving AKT1 that are corrected by antipsychotic medication. *Int. J. Neuropsychopharmacol.* **16**, 1483–1503 (2013).
- Oldham, M. C., Horvath, S. & Geschwind, D. H. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proc. Natl. Acad. Sci. USA.* **103**, 17973–17978 (2006).
- Okamura, Y. *et al.* COXPRESdb in 2015: coexpression database for animal species by DNA-microarray and RNAseq-based expression data with multiple quality assessment systems. *Nucleic Acids Res.* **43**, D82–86 (2015).
- Gaiteri, C., Ding, Y., French, B., Tseng, G. C. & Sibille, E. Beyond modules and hubs: the potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders. *Genes, brain, and behavior* **13**, 13–24 (2014).
- de la Fuente, A. From ‘differential expression’ to ‘differential networking’ - identification of dysfunctional regulatory networks in diseases. *Trends Genet.* **26**, 326–333 (2010).
- Torkamani, A., Dean, B., Schork, N. J. & Thomas, E. A. Coexpression network analysis of neural tissue reveals perturbations in developmental processes in schizophrenia. *Genome Res.* **20**, 403–412 (2010).
- Voineagu, I. *et al.* Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **474**, 380–384 (2011).
- de Jong, S. *et al.* A gene co-expression network in whole blood of schizophrenia patients is independent of antipsychotic-use and enriched for brain-expressed genes. *PLoS One* **7**, e39498 (2012).
- Ponomarev, I., Wang, S., Zhang, L., Harris, R. A. & Mayfield, R. D. Gene coexpression networks in human brain identify epigenetic modifications in alcohol dependence. *J. Neurosci.* **32**, 1884–1897 (2012).
- Chen, C. *et al.* Two gene co-expression modules differentiate psychotics and controls. *Mol. Psychiatry* **18**, 1308–1314 (2013).
- Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
- Gurling, H. M. *et al.* Genomewide genetic linkage analysis confirms the presence of susceptibility loci for schizophrenia, on chromosomes 1q32.2, 5q33.2, and 8p21–22 and provides support for linkage to schizophrenia, on chromosomes 11q23.3–24 and 20q12.1–11.23. *Am. J. Hum. Genet.* **68**, 661–673 (2001).
- Lewis, C. M. *et al.* Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *Am. J. Hum. Genet.* **73**, 34–48 (2003).
- Bly, M. Mutation in the vesicular monoamine gene, SLC18A1, associated with schizophrenia. *Schizophr. Res.* **78**, 337–338 (2005).
- Richards, M. *et al.* Association study of the vesicular monoamine transporter 1 (VMAT1) gene with schizophrenia in a Japanese population. *Behavioral and brain functions: BBF* **2**, 39 (2006).
- Chen, S. F. *et al.* Support for association of the A277C single nucleotide polymorphism in human vesicular monoamine transporter 1 gene with schizophrenia. *Schizophr. Res.* **90**, 363–365 (2007).
- Lohoff, F. W. *et al.* Association between polymorphisms in the vesicular monoamine transporter 1 gene (VMAT1/SLC18A1) on chromosome 8p and schizophrenia. *Neuropsychobiology* **57**, 55–60 (2008).
- Gough, S. C., Walker, L. S. & Sansom, D. M. CTLA4 gene polymorphism and autoimmunity. *Immunol. Rev.* **204**, 102–115 (2005).
- Muller, N. & Schwarz, M. J. Immune System and Schizophrenia. *Curr. Immunol. Rev.* **6**, 213–220 (2010).
- Liu, J. *et al.* CTLA-4 confers a risk of recurrent schizophrenia, major depressive disorder and bipolar disorder in the Chinese Han population. *Brain. Behav. Immun.* **25**, 429–433 (2011).
- Jones, A. L. *et al.* CTLA-4 single-nucleotide polymorphisms in a Caucasian population with schizophrenia. *Brain. Behav. Immun.* **23**, 347–350 (2009).
- Kordi-Tamandani, D. M., Vaziri, S., Dahmardeh, N. & Torkamanzahi, A. Evaluation of polymorphism, hypermethylation and expression pattern of CTLA4 gene in a sample of Iranian patients with schizophrenia. *Mol. Biol. Rep.* **40**, 5123–5128 (2013).
- Jun, T. Y. *et al.* Polymorphism of CTLA-4 gene at position 49 of exon 1 may be associated with schizophrenia in the Korean population. *Psychiatry Res.* **110**, 19–25 (2002).
- Merzdorf, C. S. Emerging roles for zic genes in early development. *Dev. Dyn.* **236**, 922–940 (2007).
- Ogura, H., Aruga, J. & Mikoshiba, K. Behavioral abnormalities of Zic1 and Zic2 mutant mice: implications as models for human neurological disorders. *Behav. Genet.* **31**, 317–324 (2001).
- Hatayama, M. *et al.* Zic2 hypomorphic mutant mice as a schizophrenia model and ZIC2 mutations identified in schizophrenia patients. *Sci. Rep.* **1**, 16 (2011).
- Pourebahram, R. *et al.* Transcription factor Zic2 inhibits Wnt/beta-catenin protein signaling. *J. Biol. Chem.* **286**, 37732–37740 (2011).
- Freyberg, Z., Ferrando, S. J. & Javitch, J. A. Roles of the Akt/GSK-3 and Wnt signaling pathways in schizophrenia and antipsychotic drug action. *Am. J. Psychiatry* **167**, 388–396 (2010).
- Katsel, P., Tan, W., Fam, P., Purohit, D. P. & Haroutunian, V. Cell cycle checkpoint abnormalities during dementia: A plausible association with the loss of protection against oxidative stress in Alzheimer’s disease [corrected]. *PLoS One* **8**, e68361 (2013).
- Emamian, E. S., Hall, D., Birnbaum, M. J., Karayiorgou, M. & Gogos, J. A. Convergent evidence for impaired AKT1-GSK3beta signaling in schizophrenia. *Nat. Genet.* **36**, 131–137 (2004).
- Keri, S., Beniczky, S. & Kelemen, O. Suppression of the P50 evoked response and neuregulin 1-induced AKT phosphorylation in first-episode schizophrenia. *Am. J. Psychiatry* **167**, 444–450 (2010).

42. Emamian, E. S. AKT/GSK3 signaling pathway and schizophrenia. *Front. Mol. Neurosci.* **5**, 33 (2012).
43. Ma, X., Fei, E., Fu, C., Ren, H. & Wang, G. Dysbindin-1, a schizophrenia-related protein, facilitates neurite outgrowth by promoting the transcriptional activity of p53. *Mol. Psychiatry* **16**, 1105–1116 (2011).
44. Ibanez, K., Boulosa, C., Tabares-Seisdedos, R., Baudot, A. & Valencia, A. Molecular evidence for the inverse comorbidity between central nervous system disorders and cancers detected by transcriptomic meta-analyses. *PLoS Genet.* **10**, e1004173 (2014).
45. Funk, A. J., McCullumsmith, R. E., Haroutunian, V. & Meador-Woodruff, J. H. Abnormal activity of the MAPK- and cAMP-associated signaling pathways in frontal cortical areas in postmortem brain in schizophrenia. *Neuropsychopharmacology* **37**, 896–905 (2012).
46. Turetsky, B. I., Hahn, C. G., Borgmann-Winter, K. & Moberg, P. J. Scents and nonsense: olfactory dysfunction in schizophrenia. *Schizophr. Bull.* **35**, 1117–1131 (2009).
47. Moberg, P. J. *et al.* Meta-analysis of olfactory function in schizophrenia, first-degree family members, and youths at-risk for psychosis. *Schizophr. Bull.* **40**, 50–59 (2014).
48. Harris, L. W. *et al.* Comparison of peripheral and central schizophrenia biomarker profiles. *PLoS One* **7**, e46368 (2012).
49. R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*, URL <http://www.R-project.org/> (2014).
50. Orr, M. & Liu, P. *ssize.fdr*: sample size calculations for microarray experiments. *R package version 1.2* (2015).
51. Liu, P. & Hwang, J. T. Quick calculation for sample size while controlling false discovery rate with application to microarray analysis. *Bioinformatics* **23**, 739–746 (2007).
52. Warnes, G. R. *et al.* *gplots*: Various R programming tools for plotting data. *R package version 2.12.1*, <http://CRAN.R-project.org/package=gplots> (2013).
53. Gentleman, R., Carey, V., Huber, W. & Hahne, F. *genefilter*: methods for filtering genes from high-throughput experiments. *R package version 1.48.1* (2015).
54. Varemo, L., Nielsen, J. & Nookaew, I. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Res* **41**, 4378–4391 (2013).
55. Du, P. *et al.* From disease ontology to disease-ontology lite: statistical methods to adapt a general-purpose ontology for the test of gene-ontology associations. *Bioinformatics* **25**, i63–68 (2009).
56. Feng, G., Shaw, P., Rosen, S. T., Lin, S. M. & Kibbe, W. A. Using the bioconductor GeneAnswers package to interpret gene lists. *Methods Mol. Biol.* **802**, 101–112 (2012).
57. Deng, X. *geneListPie*: Profiling a gene list into GOSlim or KEGG function pie. *R package version 1.0*, <http://CRAN.R-project.org/package=geneListPie> (2012).
58. Franceschini, A. *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* **41**, D808–815 (2013).
59. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
60. Langfelder, P., Luo, R., Oldham, M. C. & Horvath, S. Is my network module preserved and reproducible? *PLoS Comput. Biol.* **7**, e1001057 (2011).
61. Langfelder, P. & Horvath, S. Fast R Functions for Robust Correlations and Hierarchical Clustering. *Journal of statistical software* **46**, 1–17 (2012).

Acknowledgements

We sincerely thank all the subjects for their support and participation and all the medical staff involved in collecting blood samples. We thank Chunyu Liu and Simone de Jong for their help in the manuscript writing. This work was supported by the National Natural Science Foundation of China (81471364, 81271482, 81571319), Beijing Natural Science Foundation (7132182), Program for New Century Excellent Talents in University (NCET-12-1036). Shugart was supported by the Intramural Research Program of National Institute of Mental Health, National Institutes of Health (MH002929-05).

Author Contributions

Y.X. and F.Z. designed the study; F.Z., W.Y. and Y.Y.S. did the statistical analysis and wrote the manuscript; G.W., K.Z., C.J., J.W., H.Y. and D.Z. contributed to administrative and material support.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Xu, Y. *et al.* Altered expression of mRNA profiles in blood of early-onset schizophrenia. *Sci. Rep.* **6**, 16767; doi: 10.1038/srep16767 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>