Differential Expression of Long Noncoding RNAs in Peripheral Blood Leukocytes of Patients with Schizophrenia

Jie Wu
Sixth Affiliated Hospital of Kunming Medical University: People's Hospital of Yuxi City
Zijun Liu
Kunming Medical University Sixth Affiliated Hospital
Yunqiao Zhang
Kunming Medical University Sixth Affiliated Hospital
Zhaowei Teng
Kunming Medical University Sixth Affiliated Hospital
Xu You
Kunming Medical University Sixth Affiliated Hospital
Ye Ruan
Kunming Medical University Sixth Affiliated Hospital
Changqing Gao
Kunming Medical University Sixth Affiliated Hospital
Yong Zeng (✉ zengyong1388851038@163.com)
Kunming Medical University Sixth Affiliated Hospital https://orcid.org/0000-0001-8512-195X

Research article

Keywords: Schizophrenia, long non-coding RNA, RNA sequencing, bioinformatics, IL1RAP

Posted Date: March 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-113426/v2

Abstract

Background: The diagnosis of schizophrenia (SCZ) depends on the evaluation of clinical symptoms, but there is no objective biomarker. Studies have shown that long non-coding RNAs (lncRNAs) may be involved in the pathogenesis of SCZ. In this study, we evaluated the differential gene expression of lncRNAs and mRNAs in peripheral blood (PBL) of patients with SCZ.

Methods: We profiled the transcriptome of PBL in 50 patients with SCZ and 50 controls without psychiatric diagnoses, using RNA-seq. LncRNA-mRNA interactions were predicted using “RNAplex”, a hierarchical classification-Spielman correlation coefficient approach was used to analyze the correlation between lncRNAs and protein-coding gene expression among samples, and systematic bioinformatic methods (Go/Pathway) were used to perform lncRNA functional annotation. The results were validated using qPCR. Functional sites in sequences were predicted using the PROSITE, NCBI, UCSC, and JASPAR databases.

Results: We identified 94 lncRNAs and 1179 mRNAs differentially expressed in PBL, of which 46 lncRNAs were identified for the first time. Enrichment of lncRNAs involved biological processes and signaling pathways related to neutrophil activation involved in the immune response. Spearman correlation coefficient analysis showed that 81 lncRNAs and 410 mRNAs had correlated expression ($p < 0.01$ and $|r| \geq 0.4$). qPCR performed on independent samples verified that the core node of the lncRNA-mRNA co-expression network, the IL1RAP-TCONS_00138311 variable splicing transcript, was highly expressed in patients with SCZ ($2^{\triangle\triangle Ct}$ of 0.56, area under the ROC curve of 0.924). The top four ranked transcription factors were predicted to be HSF1, HSF2, HSF4, and FOXA1.

Conclusions: Combined with sequence function analysis, we showed that the transcription factors FOXA1, HSF1, HSF2, and HSF4 may mediate the activation of IL1B-induced NF-kβ pathway and other inflammatory pathways through regulation of the IL1RAP alternative splicing transcript TCONS_00138311, thereby participating in the pathogenesis of SCZ. We propose that the frequency of differentially expressed lncRNAs in PBL may serve as a novel biomarker for diagnosis of SCZ.

Background

Schizophrenia (SCZ) is a serious mental disorder of uncertain etiology, with a high incidence of morbidity, recurrence, and disability, as well as a lifetime prevalence of approximately 1% in the adult population [1]. SCZ has a high prevalence of residual symptoms, with approximately 10–15% of patients requiring lifelong care [2]. Till date, SCZ is diagnosed empirically by clinicians, and there are no objective laboratory indicators. Elucidating the etiology and pathological mechanism of SCZ, identifying clinical diagnostic biomarkers, and developing targeted drug therapy have valuable application and significance.

Advances in genome-wide association studies (GWAS) indicate that the occurrence of SCZ is affected by a combination of genetic and environmental factors [3]. Recent studies have shown that long non-coding RNAs (lncRNAs) may be involved in the pathogenesis of SCZ [4]. LncRNAs play important roles in many

life processes, such as dose compensation, epigenetic regulation, cell cycle regulation, regulation of cell differentiation, transcription activation, and transcription interference [5]. Pro-inflammatory cytokines, such as IL-6, TNF-α, and IFN-γ, have been reported to be significantly elevated in SCZ; in contrast, lncRNAs, such as TMEVPG1, enhance the expression of IFN-γ, and NRON regulates the transcription factor NFAT, thereby regulating IL-6 and IFN-γ expression. Conversely, these inflammatory cytokines can stimulate and alter lncRNA expression. LncRNAs are mostly located in the nucleus and are highly expressed in the mammalian brain [6]. As brain tissue and peripheral blood (PBL) gene expression have a common regulatory pathway [7], multiple cytokines secreted in the brain are also present in peripheral blood leukocytes. Therefore, many lncRNAs that are highly expressed in brain tissue are highly expressed in PBL as well. It is noteworthy that the detection of lncRNA expression levels has high feasibility and reliability. In this study, we analyzed the transcriptome of PBL to evaluate potential lncRNA markers of SCZ.

Methods

Study Samples

Blood samples used for sequencing were collected from patients in the psychiatric inpatient or outpatient departments of the Sixth Affiliated Hospital of Kunming Medical University, the Second People's Hospital of Honghe Prefecture, and the Second People's Hospital of Yuxi City. Patients (50 cases) in the test groups conformed to the diagnosis of SCZ according to the fourth edition of the American Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). Patients with SCZ were at the initial onset or did not take antipsychotic drugs in the five weeks before enrollment, belonged to Han ancestry, and were aged 15 to 60 years. A control group was recruited at the same conditions.

RNA Sequencing of PBL

After drawing 5 ~ 10 mL of venous blood into tubes containing the anticoagulant EDTA, centrifugation was performed within 60 min of collection. Total RNA was extracted using the TRIZOL Reagent (Ambion®), and the VAHTS Total RNA-seq (H/M/R) Library Prep Kit for Illumina® (NR603) kit was used for RNA purification and library construction. The constructed sequencing library was used for RNA sequencing (RNA-seq) in Illumina HiSeq TM 4000 (China, Illumina). Illumina HiSeq TM4000 is currently the only expression profile chip that can achieve 30 repetitions of the probe, and the chip results and qPCR show a correlation coefficient $R^2 = 0.97$ [8]. RNA-seq compared with the other sequencing techniques has a broader dynamic range, provides a better estimate of relative expression levels of any genomic region with higher technical reproducibility, and facilitates the detection of alternative splicing [9]. With these advantages, RNA-seq has been used to reconstruct the entire transcriptome of an organism [10]. The novel lncRNA predictions for the selected new transcripts, for the coding capacity of new transcripts using CPC, CNCI, and SwissProt, and for the intersection of these transcripts without coding potential and protein annotation information were considered reliable.

Gene Co-Expression Network Construction

Differential analysis of lncRNA expression between groups was performed using edgeR [11]. Differential transcripts were selected using FDR and log2FC, and the filtering conditions were FDR < 0.05 and |log2FC| > 1. As lncRNA expression is low compared to mRNA levels, a hierarchical classification-Spielman correlation coefficient approach was used to analyze the correlation between lncRNA and protein-coding gene expression among samples. The Cytoscape software was used to build the co-expression network.

LncRNA Action Modes

The action modes of lncRNAs involved in regulation are divided into three categories: antisense, cis, and trans [12]. We predicted target genes by the modes of action, in which each lncRNA is likely to be involved in the following manner: Antisense action target gene prediction was performed using RNAplex to find short interactions between two long-stranded RNAs and to find complementary binding between antisense lncRNAs and mRNAs. Cis action target gene prediction was performed by annotating lncRNAs within 10 kb upstream or downstream of a gene, and analyzing their potential to intersect with regions where cis-acting elements are located. Trans-action target gene prediction was performed using Pearson's correlation coefficient method to analyze the correlation of lncRNAs and protein-coding gene expression between samples, and protein-coding genes with absolute correlation values greater than 0.999 were selected.

Targeted mRNA Enrichment Analysis (Go/Pathway)

Using the DAVID database (V.6.8) and the "clusterProfiler" R package [13], Gene Ontology (GO) enrichment analysis of 410 differentially expressed mRNAs was performed to annotate the function of lncRNAs [14]. Signaling pathways were also annotated using the Kyoto Encyclopedia of Genes and Genomes (KEGG) [15].

Quantitative Polymerase Chain Reaction (qPCR)-based Validation of LncRNAs

The selected core lncRNAs were subjected to qPCR analysis for validation. The filtering conditions were FDR < 0.05, |log2FC| > 1, FPKM > 0.5, and GO/Pathway enrichment analysis of functional lncRNAs. β-Actin was used as an internal reference gene to compare the control group (CK) and test group (T), using the $2^{\triangle\triangle Ct}$ method. The lncRNA expression rates between the CK and T groups were compared using the independent samples t-test. ROC curve statistical test was used to determine whether lncRNAs can be used as a possible diagnostic biomarker for SCZ. The SPSS 22.0 software was used to analyze and process the data, and statistically significant differences are indicated by $p < 0.05$.

Functional Analysis of qPCR-validated LncRNAs

Functional site prediction in qPCR-validated transcripts was performed using the PROSITE database [16]. If the candidate sequence was found to have no transcription or translation function or was not itself a transcription factor, then an NCBI blast search was performed to predict whether the lncRNAs are transcripts of target genes and their homologues [17]. A conservative analysis was then performed via the UCSC database [18], and transcription factors that bind to lncRNAs were searched. Finally, the

JASPAR database was used to predict transcription factors and binding sites [19], selecting transcription factors with the same direction of transcription.

Results

Clinical Characteristics of Patients that Provided Sequencing Samples

Statistical differences were evaluated in sex and age between the CK (n = 50) and T (n = 50) groups. Analysis using the independent samples t-test showed that the mean age ± standard deviation of the T and CK groups were 35.56 ± 10.04 and 32.52 ± 8.04, respectively, and had no significant difference ($p > 0.05$). There was no significant difference in gender composition ($p > 0.05$).

RNA-seq Quality Assessment

RNA quality of the sequencing samples was analyzed using an Agilent 2100 Bioanalyzer. The OD_{260}/OD_{280} values were all between 1.8-2.0 (Fig. 1a). The two-sample independent repeated correlation is shown in Fig. 1b, which shows that the quality of the isolated RNA met the requirements of sequencing. The Euclidean distance algorithm was used to calculate the expression distance of each sample gene, the Pearson algorithm was used to calculate the distance between the samples, and then a cluster map was established according to the distance (Fig. 1c). A total of 5106 new lncRNA transcripts were obtained by CPC, CNCI, and SwissProt prediction (Fig. 1d). Principal component analysis of sequencing samples showed that the PC1 principal component was 79.4%, PC2 was 9.3%, and that PC1 + PC2 could distinguish 88.7% of the overall variance (Fig. 1e).

Gene Co-Expression Network Construction

A total of 94 differentially expressed lncRNAs and mRNAs with FDR < 0.05 and |log2FC| > 1 were identified (Fig. 2a, 2c), of which 70 were upregulated and 24 downregulated. Forty-six new lncRNAs were identified for the first time, of which XLOC_079942, LINC02362, ACTA2-AS1, AL162431.2, and CLSTN2-AS1 were the top five upregulated lncRNAs (|log2FC| > 5), and AL390763.2, XLOC_078139, XLOC_043880, AC008892.1, and AL034397.3 were the top five downregulated ones (|log2FC| > 4).

A total of 1179 differentially expressed mRNAs were obtained (Fig. 2b, 2d), of which 702 were upregulated and 477 downregulated. CNOT2, RNF19B, MECP2, TNFAIP3, and ALDO were the top five upregulated mRNAs (|log2FC| > 9) and AHR, ARMC8, HBP1, ERVFC1-1, and COTL1 were the top five downregulated ones (|log2FC| > 7). Heat maps of the expression patterns of 94 lncRNAs and 1179 mRNAs showed distinct group information after clustering between groups. The SCZ group was clearly distinguished from the normal group, suggesting that the differentially expressed genes may serve as potential biomarkers of SCZ.

According to Spearman correlation coefficient analysis, 81 lncRNAs and 410 mRNAs showed correlated expression ($p < 0.01$ and $|r| \geq 0.4$), 1152 mRNAs were positively correlated with lncRNAs, and 300 mRNAs

were negatively correlated with lncRNAs (Fig. 3).

The obtained data and Cytoscape software were used to construct a lncRNA-mRNA co-expression network, and independent individuals were excluded.
Some core lncRNAs, such as PDCD1, MEG3, LINC00599, PSMA3-AS1, CLSTN2-AS1, AC008892.1, and XLOC078139, which target multiple mRNAs, were identified in the co-expression network. The core node IL1RAP was also found to be linked to several genes, suggesting that it could be a candidate gene for late functional validation.

LncRNA Action Modes

Target gene prediction of lncRNA mode of action yielded one lncRNA with an antisense interaction with mRNA, 22 lncRNAs with cis action, and 94 trans-acting lncRNAs (Fig. 4). Among them, XLOC_079942, XLOC_081142, LINC00958, LINC00599, LINC01052, CLSTN2-AS1, AL034397.3, and LINC01287 had more trans-acting association with mRNAs as members of the core node. IL1RAP was associated with multiple lncRNAs. The results showed that lncRNAs mainly affected mRNA expression in a co-expression manner, and thus participated in gene expression regulation and transcription.

Results of GO/Pathway Enrichment Analysis

GO/Pathway enrichment analysis of 410 target mRNAs was performed and filtered at $p < 0.05$. A total of 271 mRNAs were enriched in biological process (BP) terms, with the first four items being significantly related to neutrophil activation, neutrophil degranulation, neutrophil activation involved in immune response, neutrophil-mediated immunity, and leukocyte migration (all related to immune response) (Fig. 5a, 5b). Twenty-three mRNAs were enriched in the following cellular component terms: specific granule lumen, secret granule lumen, and cytoplasmic vesicle lumen (Fig. 5c, 5d). A total of 16 mRNAs were enriched in molecular function (MF) terms, including protein binding, transcription factor binding, and transcription regulatory region DNA binding (Fig. 5e, 5f). The first four of these BPs share the genes CAPN1, PLAU, OLR1, CTSD, ABCA13, LCN2, MS4A3, and TCN1, and are all involved in the pathogenesis of neuropsychiatric disorders.

The KEGG enrichment results also showed that the main enrichment pathways for targeted mRNAs were concentrated in immune response, such as C-type lectin receptor, NOD-like receptor, TNF, and the NF-kappa β signaling pathways (Fig. 6a–6d).

qPCR Validation Results

The mean Ct values of β-actin obtained in PBL samples from CK and T groups were 15.44–15.24, and the CV% was 3.38–3.05.
We performed qPCR experiments with the lncRNAs selected by bioinformatic analysis to meet the conditions of FDR < 0.05, |log2FC| > 1, and differential expression of FPKM > 0.5 between groups. The experimental results were as follows: IL1RAP-TCONS_00138311 was downregulated 0.56-fold in the T group relative to the CK group ($p < 0.05$) (Fig. 7a, Table 1), and CCR3-TCONS_00134168 was upregulated

4.68-fold in the T group relative to the CK group ($p < 0.05$) (Fig. 7b, Table 1). The sequencing results were considered accurate and credible, and were subjected to further analysis and functional verification.

Taking IL1RAP-TCONS_00138311 and CCR3-TCONS_00134168 as test variables, and the CK and T groups as static variables, a ROC curve was constructed. The results showed that the expression of IL1RAP-TCONS_00138311 and CCR3-TCONS_00134168 in the PBL leukocytes of patients with SCZ had certain effects on the predicted value ($p < 0.05$). Using IL1RAP-TCONS_00138311 as a diagnostic indicator, AUC = 0.924, 95% of the confidence interval was 0.873–0.974, and the Jorden index reached its maximum value of 0.74 when FPKM = 15.48 in PBL, at which point the sensitivity was 80.0% and the specificity 94.0% (Fig. 7c, Table 1). Principal component analysis, using IL1RAP-TCONS_00138311 sequencing expression value, showed that the PCA1 + PCA2 components explained 24.3% of the overall variance (Fig. 7d).

Table 1
qPCR and ROC analyses results

LncRNA	IL1RAP-TCONS_00138311	CCR3-TCONS_00134168
Sequence (5' to 3')	F:AGGGGAAGGGAATCAACAAATAG	F:TCAAGACTTCGTGGCTTAAACAATA
	R:GGGGCGTGGCATGTAACC	R:GGAACTCCATACCTGAAAGACCCTA
CK	7.13 ± 0.69	14.59 ± 0.88
T	8.58 ± 1.06	12.93 ± 1.43
$2^{\triangle\triangle Ct}$*	0.56	4.68
AUC	0.9236	0.6486
P-value*	0.009***	0.016**
*△△Ct = Test group△Ct-Control group△Ct		
*$2^{\triangle\triangle Ct}$ = Expression ratio, $2^{\triangle\triangle Ct} > 1$ = upregulated, $2^{\triangle\triangle Ct} < 1$ = downregulated.		

IL1RAP-TCONS_00138311 Sequence Functional Analysis

The selected transcripts were predicted using the CPC, CNCI, and SwissProt databases, and none had coding potential or protein annotation information. The IL1RAP-TCONS_00138311 transcript sequence was used for an NCBI blast search, with Query Length 1311, bits to 1023 length, and matches to Sequence ID: NG_029105.2 (43715–44737, Score = 1890, Expect = 0.0, Identities = 100%, Gaps = 0%), and was identified as a new IL1RAP transcript. Using the upstream 2000 bp and the downstream 100 bp sequences as the gene promoter region, a conservative analysis through the UCSC database identified transcription factors that can bind to the lncRNA (Fig. 8a). Finally, using the JASPAR database, with score set to 500 and relative profile score threshold set at 90%, the transcription factors in the same transcription direction as IL1RAP-TCONS_00138311 were selected. The top four ranked transcription factors were predicted to be HSF1, HSF2, HSF4, and FOXA1, and the binding sites are shown in Fig. 8b.

Discussion

Our results show that lncRNAs mainly affect mRNA expression in a co-expression manner and thus participate in gene expression regulation and transcription. IL1RAP, identified as a core node, was also found to be functionally linked to several genes. The differential expression of the variable splicing transcript TCONS_00138311 of IL1RAP was verified using qPCR, and a higher ROC area of 0.924 could specifically distinguish between patients with SCZ and healthy individuals.

The GO/Pathway analyses showed 271 transcripts enriched in BP terms. Among them, the first four of these BPs shared the genes CAPN1, PLAU, OLR1, CTSD, ABCA13, LCN2, MS4A3, and TCN1, with all of them involved in the pathogenesis of neuropsychiatric disorders. OLR1, CTSD, and MS4A3 are involved in the occurrence and development of neurodegenerative diseases. Interaction between OLR1 and the LDL component, L5, inhibits the nerve growth factor-induced activation of the PI3k/Akt cascade and suppresses cell viability and neuronal differentiation induced by nerve growth factor [20]. CTSD is expressed at high levels in almost all tissues of the brain and is essential for maintaining a steady state of lysosome-dependent proteins in this organ [21]. Defective or reduced activity of CTSD expression can lead to pathological protein aggregation in Parkinson's disease and Alzheimer’s disease (AD) [22]. The MS4A3 gene encodes a member of the transmembrane 4A gene family that exhibits a unique expression pattern in hematopoietic cells and non-lymphoid tissues, which may play a role in signal transduction and may function as a subunit associated with receptor complexes. Polymeric SNP association analysis of this gene showed a correlation with AD genes and cognitive decline [23]. CAPN1, PLAU, and TCN1 are involved in neuroprotection. CAPN1 calpain is a calcium-dependent protease, which is related to synaptic plasticity and neuroprotection in the mammalian brain [24] and regulates Alzheimer's disease-related protein signaling pathways [25]. PLAU is a urokinase-type plasminogen activator ligand that promotes neuroprotection with the receptor PLAUR or its plasma membrane-binding partner platelet-derived growth factor receptor beta (PDGFRβ) [26]. The TCN1 gene encodes a member of the vitamin b12-binding protein family, which is a major component of secondary granules in neutrophils. Vitamin B12 deficiency disrupts neurodevelopment during pregnancy and throughout childhood [27], and adequate levels of this vitamin are also necessary for adult neurocognitive function [28]. Elevated TCN1 expression is significantly associated with poorer memory [29]. ABCA13 and LCN2 are significantly associated with schizophrenia. The ABCA13 gene encodes a transmembrane transport protein, with a C-terminal motif involved in lipid transport in lipid metabolism [30], and dysregulation of lipid metabolism is strongly associated with psychiatric disorders, with ABCA13 increasing the susceptibility to SCZ, bipolar disorder, and major depression [31]. LCN2 plays an important role in mediating the inflammatory activity by binding to chemotactic peptides, leukotrienes, and platelet-activating factors [32], and is significantly increased in SCZ [33]. Proteins, such as TNIP1 and IL1B, regulate NF-kβ activation and play a role in autoimmunity and tissue homeostasis; MGST1, CAMP, TNFRSF14, and FOS are involved in cell chemotaxis, immune mediator induction, and inflammatory response regulation.

The KEGG enrichment results also showed that the main enriched pathways for targeted mRNAs were concentrated in the immune response, such as C-type lectin receptor, NOD-like receptor, TNF, and the NF-

kappa β signaling pathways (Fig. 6a-6d). The nuclear factor-kβ (NF-kβ) complex provides direct transcriptional control of interferon-induced transmembrane protein (IFITM) and cytokines [34–36], which are overexpressed in SCZ [37]. In the prefrontal cortex (PFC) of schizophrenic subjects, the mRNA levels of two NF-kβ family members, NF-kβ1 and NF-kβ2, increased, whereas the mRNA levels of NF-kβ site-binding protein Schnurri, which inhibits the function of NF-kβ-2, were low [36]. These findings suggest that higher NF-kβ transcriptional activity may play an important role in elevating IFITM and cytokine mRNA levels in the PFC of patients with SCZ.

IL1RAP is an essential protein for the IL1 receptor [38], regulating the IL1B-induced activation of the NF-kβ pathway and other inflammatory pathways, which is closely related to SCZ. In addition, the IL1B inflammatory pathway plays an important role in the development of SCZ [39]. There are currently four main types of IL1RAP transcripts and protein isomers. Among them, mIL1RAP and sIL1RAP are the main types, with different molecular structures, tissue distribution, and biological functions [40, 41]. mIL1RAP has extracellular, transmembrane, and intracellular segments, which affect IL1B downstream signaling and activating transcription and translation of IL1B effector genes [42]. sIL1RAP only has the extracellular segment that can bind to IL1B, as absence of an intracellular segment prevents activation of the IL1B pathway.

Conclusions

In this RNA-seq study, we used more samples for whole transcriptome sequencing and screened the differentially expressed lncRNAs, cicrRrna, microRNAs, and small RNAs in the PBL of patients with SCZ. Among them, the variable splicing transcript TCONS_00138311 of IL1RAP caught our attention. IL1RAP is an essential protein for the IL1 receptor, regulating the IL1B-induced activation of the NF-kβ pathway and other inflammatory pathways, which is closely related to SCZ. The limitation of this research is that the relevant discussions only focus on the bioinformatics analysis stage. Although the differential expression of the target gene is verified at the molecular level, there is still a lack of systematic molecular biology and zoology experimental verification; but the relevant research results are also provides precise directions for subsequent experiments, and we are also conducting cell-level experiments based on the results of this research. Hence, the next step we aimed to investigate the regulation of IL1RAP variable splicing by the transcription factors FOXA1, HSF1, HSF2, and HSF4, to clarify whether IL1RAP aberrant variable splicing is involved in the pathogenesis of SCZ, or whether the IL1RAP-TCONS_00138311 new transcripts induce the activation of NF-kβ pathway and other inflammatory pathways through IL1B, and thus participate in the development of SCZ (Fig. 8c).

Abbreviations

SCZ
schizophrenia; PBL:peripheral blood; GWAS:genome-wide association study; lncRNA:long non-coding RNA; IL1RAP:interleukin 1 receptor accessory protein; HSF1:heat shock transcription factor 1; HSF2:heat shock transcription factor 2; HSF4:heat shock transcription factor 4; FOXA1:Forkhead box A1.

Declarations

Ethics Approval and Consent to Participate

IRB approval was obtained from the Sixth Affiliated Hospital of Kunming Medical University, the Second People's Hospital of Honghe Prefecture, and the Second People's Hospital of Yuxi City. Voluntary written informed consent was provided by all participants (informed consent was obtained from a parent or guardian for participants under 16 years old).

Consent for Publication

Not applicable.

Availability of Data and Materials

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Competing Interests

The authors declare that they have no competing interests.

Funding

This study was supported by the National Natural Science Foundation of China (81760253, 81960254) and the Applied Basic Research Foundation of Yunnan Province (CN) 2018FE001-(009). The funding agencies had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

YZ conceived and supervised the project. JW designed and implemented the methods, and conducted the experiments. ZJL, YQZ, ZWT, XY, YR, and CQG contributed to data acquisition. JW and ZJL wrote the manuscript. All authors have approved the manuscript.

Acknowledgments

We thank Wenyu Zhang and Ziqiao Feng (Kunming Medical University Sixth Affiliated Hospital) for their assistance in recruiting subjects for participation in this study, and in sequencing the corresponding samples.

References

1. Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia. PLoS Med. 2005;2(5):413–33.

2. Shaloo G, Gina I, Kevin J, Kristel VI. Productivity loss and resource utilization, and associated indirect and direct costs in individuals providing care for adults with schizophrenia in the EU5. Clinicoeconomics Outcomes Res Ceor. 2015;7(11):593–602.
3. Tim BB, Stephan R, Silviualin B, Sang HL, Naomi RW, Pablo VG, et al. Genome-wide association study reveals greater polygenic loading for schizophrenia in cases with a family history of illness. Am J Med Genet B Neuropsychiatr Genet. 2016;171(2):276–89.
4. Shengdong C, Xinyang S, Wei N, Lingming K, Mingjun H, Wanshuai L, et al. Aberrant expression of long non-cding RNAs in schizophrenia patients. Int Med J Exp Clin Res. 2016;22(9):3340–51.
5. Jie S, Zh M, Zhitao M, Dapeng H, Zhenzhen W, Chuanxing L, et al. Systematic analysis of genomic organization and structure of long non-coding RNAs in the human genome. Febs Lett. 2013;587(7):976–82.
6. Jinneng C. The functional role of long non-coding RNAs and epigenetics. Biol Proced Online. 2014;16(1):11.
7. Christopher HW, Akul S, Josué PS, Stephen JG, Ming TT. The utility of gene expression in blood cells for diagnosing neuropsychiatric disorders. Int Rev Neurobiol. 2011;101:41–63.
8. Sol AJ, Jong LP, Jonghwan K, Jeong HK, Yong SK, Jin CK, et al. Comparison of the MGISEQ-2000 and Illumina HiSeq 4000 sequencing platforms for RNA sequencing. Genomics Inform. 2019;17(3):32–8.
9. John CM, Christopher EM, Shrikant MM, Matthew S, Yoav G. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res. 2008;18(9):1509–17.
10. Tnicolai S, Kapila G, George AMC, Torsten O. Gene expression in *Trypanosoma brucei*: lessons from high-throughput RNA sequencing. Trends Parasitol. 2011;27(10):434–41.
11. Robinson MD, Mccarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Biogeosciences. 2010;26:139–40.
12. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell. 2011;43(6):904–14.
13. Guangchuang Y, Ligen W, Yanyan H, Qingyu H. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics A J Integr Biol. 2012;16(5):284–87.
14. Gene OC, Mulder N. The Gene Ontology (GO) project in 2006. Nucleic Acids Res. 2005;34:322–46.
15. Minoru K, Susumu G. KEGG: Kyoto Encyclopedia of Genes and Genomes[J]. Nucleic Acids Res. 2000;28(1):27–30.
16. Christian JA, Lorenzo C, Edouard de C, Petra SL, Virginie B, Amos B, et al. PROSITE, a protein domain database for functional characterization and annotation. Nucleic Acids Res. 2010;38:161–66.
17. Nuala AL, Mathew WW, Rodney J, Stacy B, Diana C, Rich H. MV, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44(D1):733–45.
18. Laurence RM, Ann SZ, Angie SH, Donna K, Robert MK, Matthew W, et al. The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res. 2013;41:64–9.

19. Oriol F, Jaime ACM, Aziz K, Robin van der L, Xi Z, Phillip AR, et al. JASPAR 2020: Update of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 2020;48(1):87–92.
20. Jizyuh W, Chioulian L, Chingtien L, Chenyen L. Electronegative low-density lipoprotein L5 impairs viability and NGF-induced neuronal differentiation of PC12 cells via LOX-1. Int J Mol Sci. 2017;18(8):1744–89.
21. Tokiko S, Yoshinori H, Ryo Y, Jun Y, Jian Z, Kiyoharu U, et al. Proteolytic degradation of glutamate decarboxylase mediates disinhibition of hippocampal CA3 pyramidal cells in cathepsin D-deficient mice. J Neurochem. 2005;94(3):680–90.
22. Eija S, Sanna P, Petter S, Aleksi H, Matti H, Jan M, et al. Cathepsin D deficiency underlies congenital human neuronal ceroid-lipofuscinosis. Brain. 2006;129(6):1438–45.
23. Jasmine N, Gregory T, Daniel SE, Jennifer SY, Kristine Y. Gene-based aggregate SNP associations between candidate AD genes and cognitive decline. Age. 2016;38(2):41–51.
24. Soomaayeh H, Kahung L, Yan L, Yubin W, Michael RF, Xiaoning B, et al. Impaired cerebellar plasticity and eye-blink conditioning in calpain-1 knock-out mice. Neurobiol Learning Mem. 2020;170:106995.
25. Wenyue S, Qian Z, Yubin W, Athar C, Qingshun QL, Sujay D, et al. Deletion of the Capn1 gene results in alterations in signaling pathways related to Alzheimer's disease, protein quality control and synaptic plasticity in mouse brain. Front Genet. 2020;334(11):1–11.
26. Smith HW, Marshall CJ. Regulation of cell signalling by uPAR[J]. Nature Rev Mol Cell Biol. 2010;11(1):23–36.
27. Dror DK, Allen LH. Effect of vitamin B12 deficiency on neurodevelopment in infants: current knowledge and possible mechanisms. Nutrition Rev. 2008;66(5):250–5.
28. Eileen M, Alastair M, David A, Ross C, Kerrie S, David W, et al. Cognitive impairment and vitamin B12: a review. Int Psychogeriatr. 2012;24(4):541–56.
29. Ibrahim AA, Torill U, Ole AA, Hansrichard B, Vidar MS, Timothy H, et al. Expression of TCN1 in blood is negatively associated with verbal declarative memory performance. Sci Rep. 2018;8(1):12654–65.
30. Michael LF, Keiichiro O, Glenn FS, Jennifer JM, Susan AB, Mason WF. ATP-binding cassette transporter A1 contains a novel C-terminal VFVNFA motif that is required for its cholesterol efflux and ApoA-I binding activities. J Biol Chem. 2004;279(46):48477–85.
31. Helen MK, Benjamin SP, Alan M, Mary PM, Dinesh CS, Allan FM, et al. A Cytogenetic abnormality and rare coding variants identify ABCA13 as a candidate gene in schizophrenia, bipolar disorder, and depression. Am J Hum Genet. 2009;85(6):833–46.
32. Nils EM, Mads H, Kaj AJ, Jesper MH, Claus B, Bo FR, et al. Plasma neutrophil gelatinase associated lipocalin (NGAL) is associated with kidney function in uraemic patients before and after kidney transplantation. BMC Nephrol. 2012;13:2–8.
33. Li W, Yupeng D, Wei W, Xuyan F, Qi X. Elevation of plasma neutrophil gelatinase-associated lipocalin (NGAL) levels in schizophrenia patients. J Affective Dis. 2018;226:307–12.

34. Laura T, Anil KT, Jason B, Marta M, Guido F. The nuclear factor kappa B signaling pathway: integrating metabolism with inflammation. Trends Cell Biol. 2012;22(11):557–66.
35. Maria TD, Jorge M. The atypical PKCs in inflammation: NF-κB and beyond. Immunol Rev. 2012;246(1):154–67.
36. Keizo T, Katsunori K, Hideo H, Koji O, Hirotaka S, Satoko H, et al. Deficiency of Schnurri-2, an MHC enhancer binding protein, induces mild chronic inflammation in the brain and confers molecular, neuronal, and behavioral phenotypes related to schizophrenia. Neuropsychopharmacology. 2013;38(8):1409–25.
37. David WV, Anjani C, Jessica RE, Kaitlyn MR, Annie EM, David AL. Molecular mechanisms and timing of cortical immune activation in schizophrenia. Am J Psychiatry. 2015;172(11):1112–21.
38. Layé S, Liège S, Li KS, Moze E, Neveu PJ. Physiological significance of the interleukin 1 receptor accessory protein. Neuroimmunomodulation. 2001;9(4):225–30.
39. Mostaid MdS, Dimitrakopoulos S, Cassandra W, Vanessa C, Cynthia SW, Everall IP, et al. An Interleukin-1 beta (IL1B) haplotype linked with psychosis transition is associated with IL1B gene expression and brain structure. Schizophr Res. 2019;204:201–5.
40. Dirk ES, Roberta H, Della F, Heather M, Hongbo C, Ann MF, et al. The soluble form of IL-1 receptor accessory protein enhances the ability of soluble type II IL-1 receptor to inhibit IL-1 action. Immunity. 2003;18(1):87–96.
41. Liselotte EJ, Alexander SW. Expression of alternatively spliced interleukin-1 receptor accessory protein mRNAs is differentially regulated during inflammation and apoptosis. Cell Signal. 2003;15(8):793–802.
42. Wesche H, Korherr C, Kracht M, Falk W, Resch K, Martin MU. The interleukin-1 receptor accessory protein (IL-1RAcP) is essential for IL-1-induced activation of interleukin-1 receptor-associated kinase (IRAK) and stress-activated protein kinases (SAP kinases). J Biol Chem. 1997;272(12):7727–31.

Figures

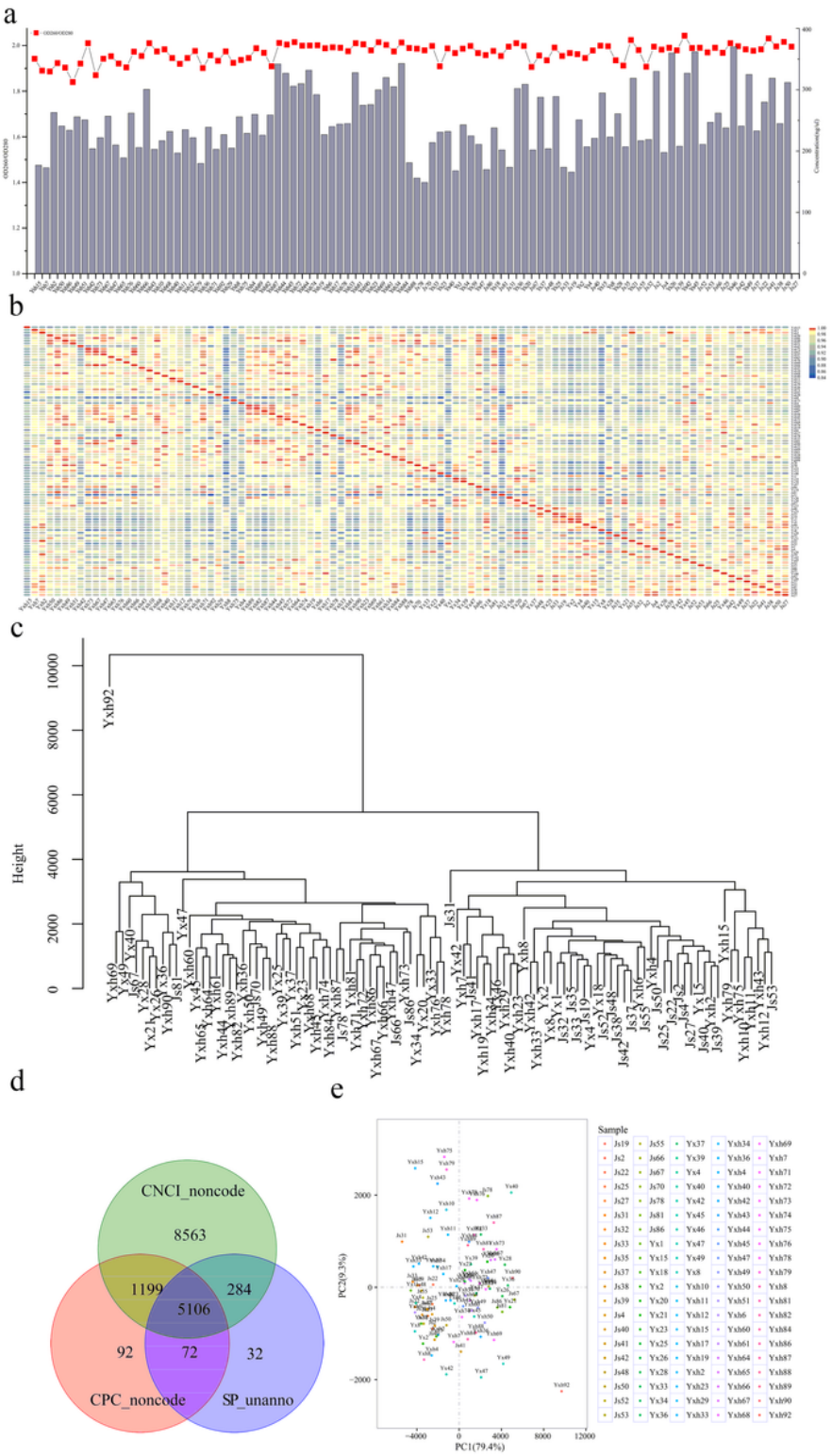

Figure 1

(a) Total RNA quality map of sequencing samples. (b) Correlation scatter plot of sequencing samples. (c) Sample clustering map. (d) Venn diagram of the new lncRNA transcripts; the intersection was predicted using CPC, CNCI, and SwissProt (5106 new lncRNAs). (e) PCA plots of sequencing samples: the PC1 and PC1+PC2 components can explain 79.4% and 88.7%, respectively, of the overall variance of the expression of all transcripts.

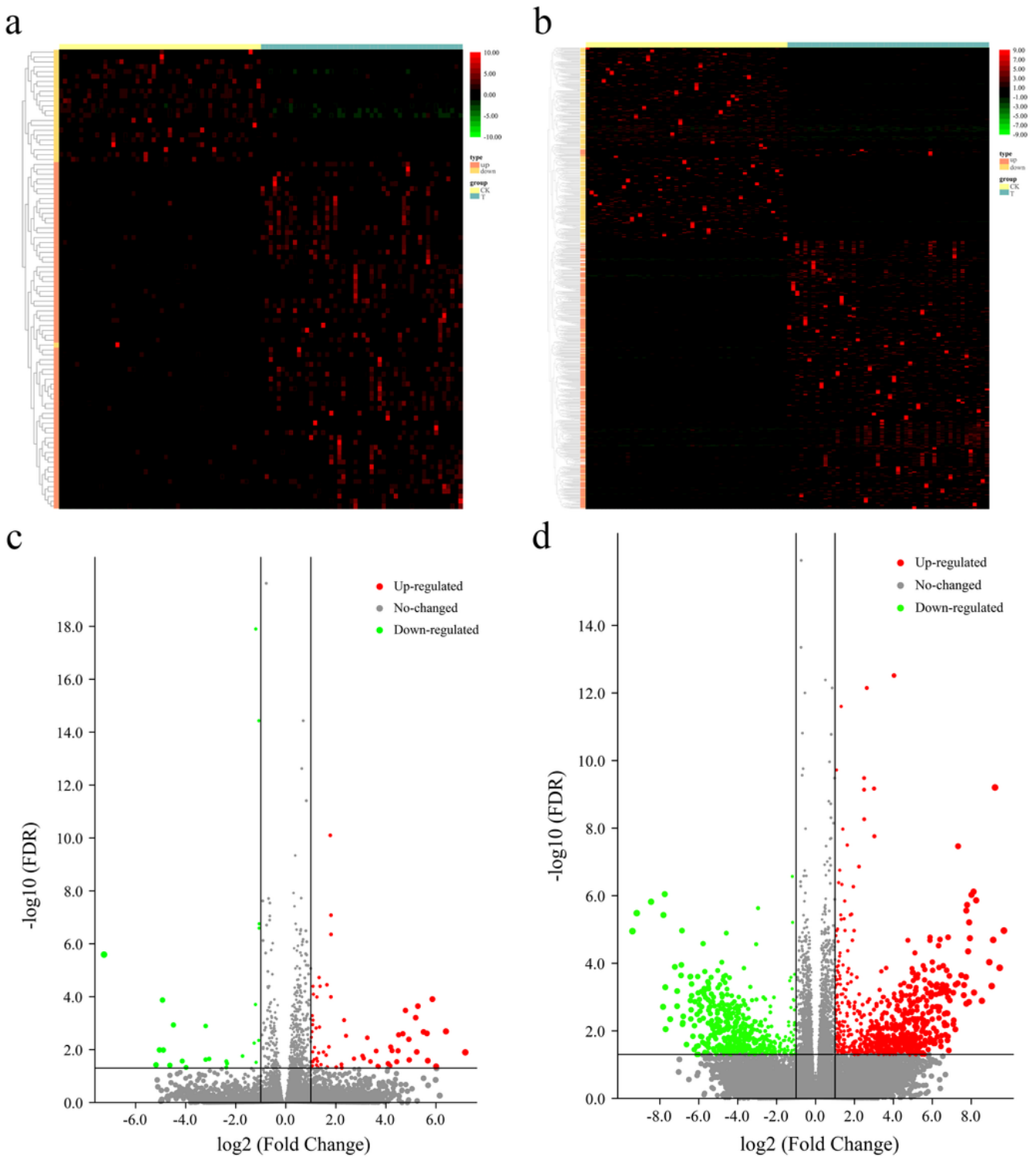

Figure 2

(a) Heatmap of expression profiles of 94 lncRNAs. (b) Heatmap of expression profiles of 1179 mRNAs. (c) Volcano plot of the FDR values as a function of weighted fold-change for lncRNAs. (d) Volcano plot of the FDR values as a function of weighted fold-change for mRNAs.

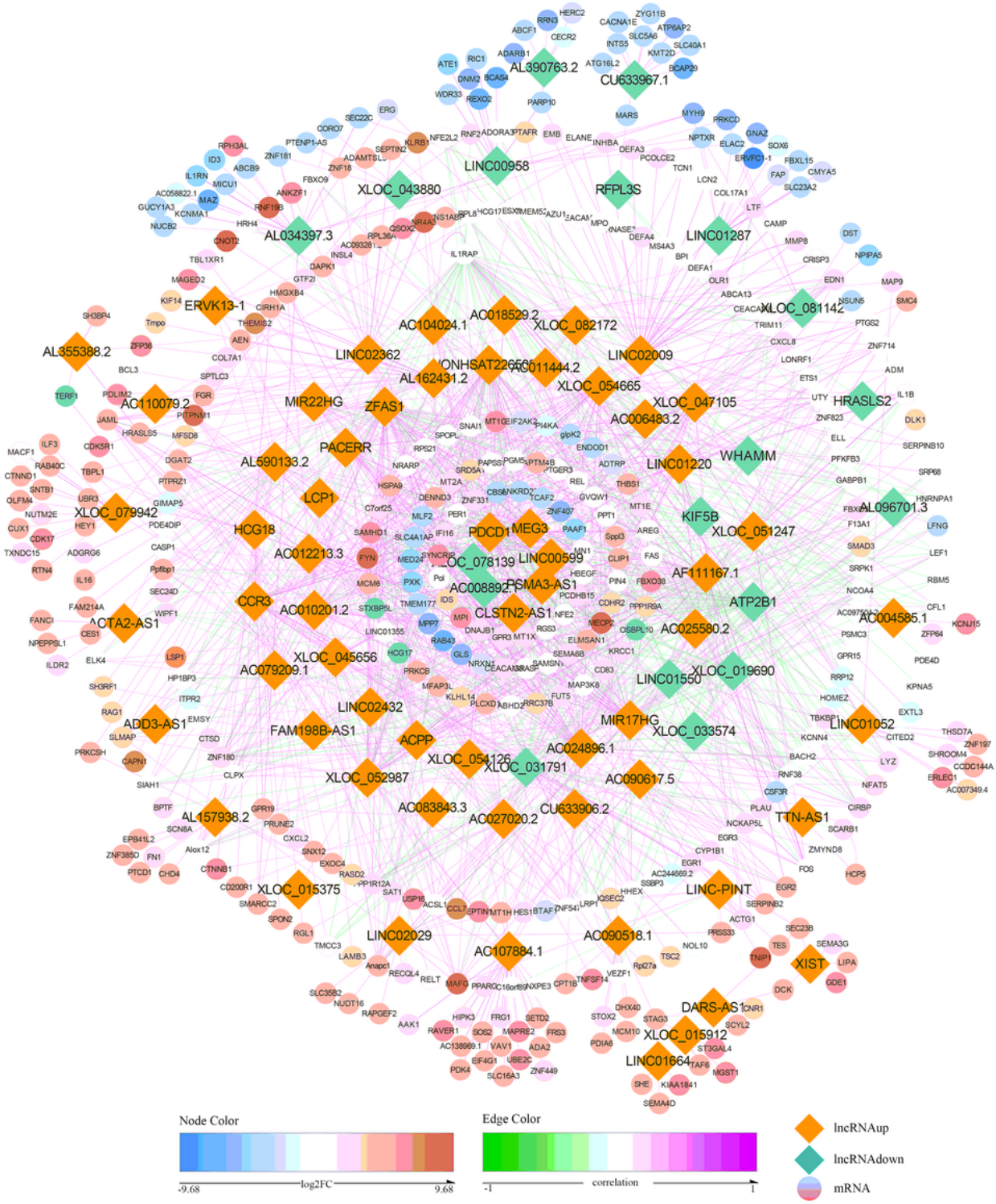

Figure 3

Predicted lncRNA and mRNA co-expression network in schizophrenia. The co-expression network was established between the 81 significantly differentially expressed lncRNAs and 410 significantly differentially expressed mRNAs that had Spearman correlation coefficients ($p < 0.01$ and $|r| \geq 0.4$). Red through blue colors indicate high to low expression levels. Purple through green colors indicate high to low correlation levels.

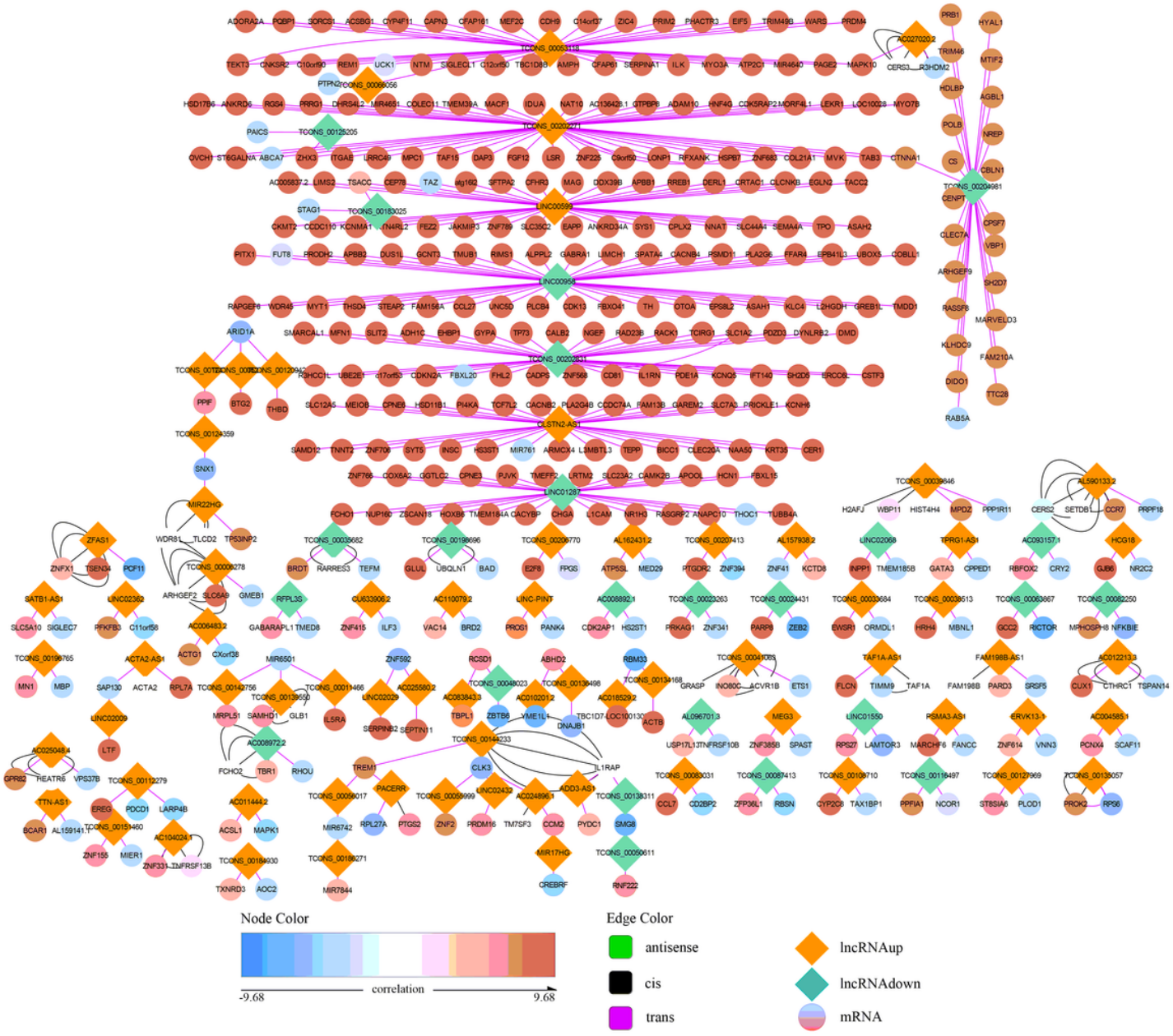

Figure 4

LncRNA-mRNA mode of action diagram. Diamonds and circles represent lncRNAs and mRNAs, respectively (colors from blue to red represent correlation size from -1 to 1); green, black, and purple lines indicate antisense, cis, and trans modes of action, respectively.

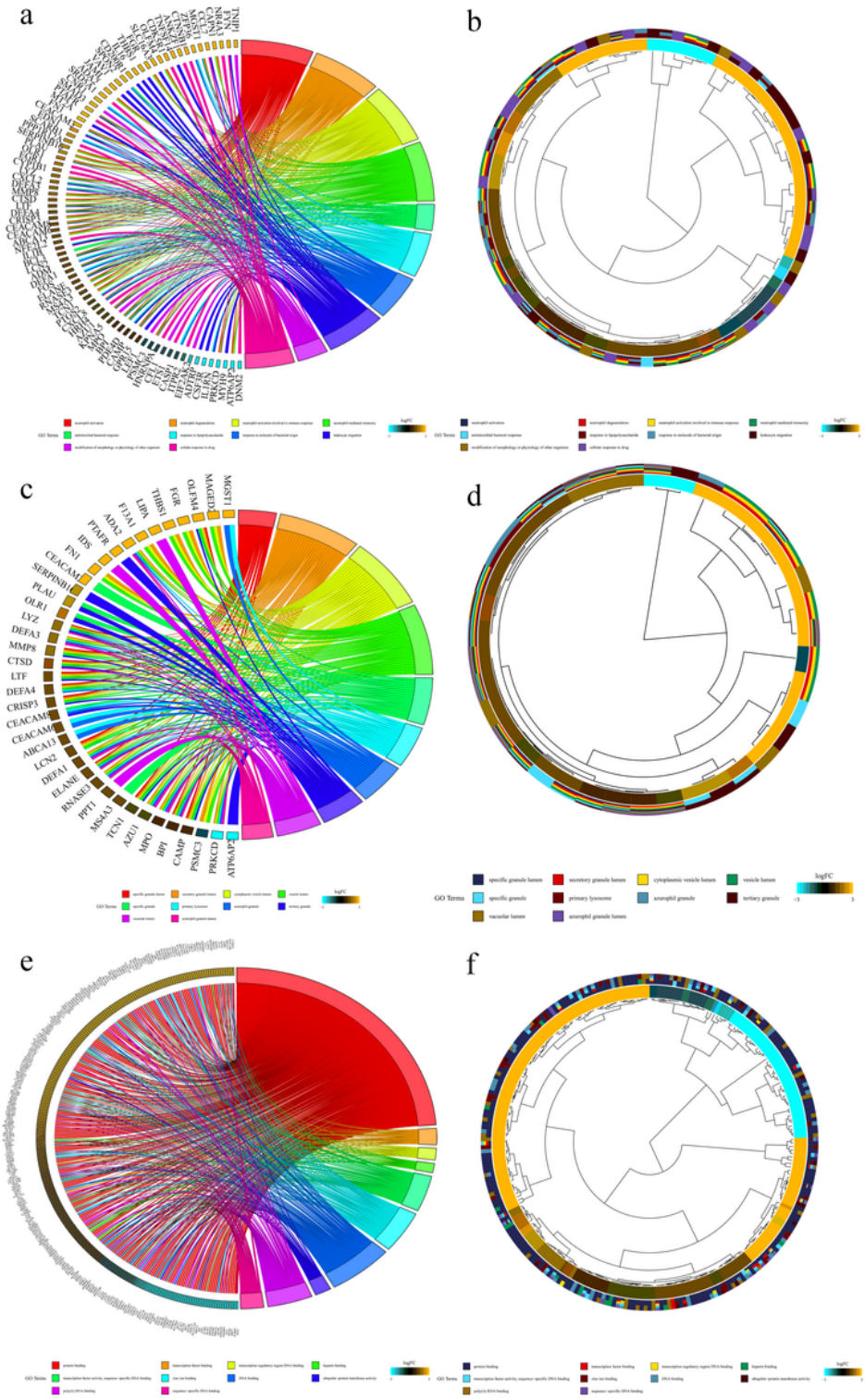

Figure 5

Top 10 GO analysis of significantly correlated mRNAs targeted by lncRNAs. (a) Chord plot display of the relationship between gene and BP_term. (b) Hierarchical clustering of gene expression profiles in each BP_term. (c) Chord plot display of the relationship between gene and CC_term. (d) Hierarchical clustering of gene expression profiles in each CC_term. (e) Chord plot display of the relationship between gene and MF_term. (f) Hierarchical clustering of gene expression profiles in each MF_term.

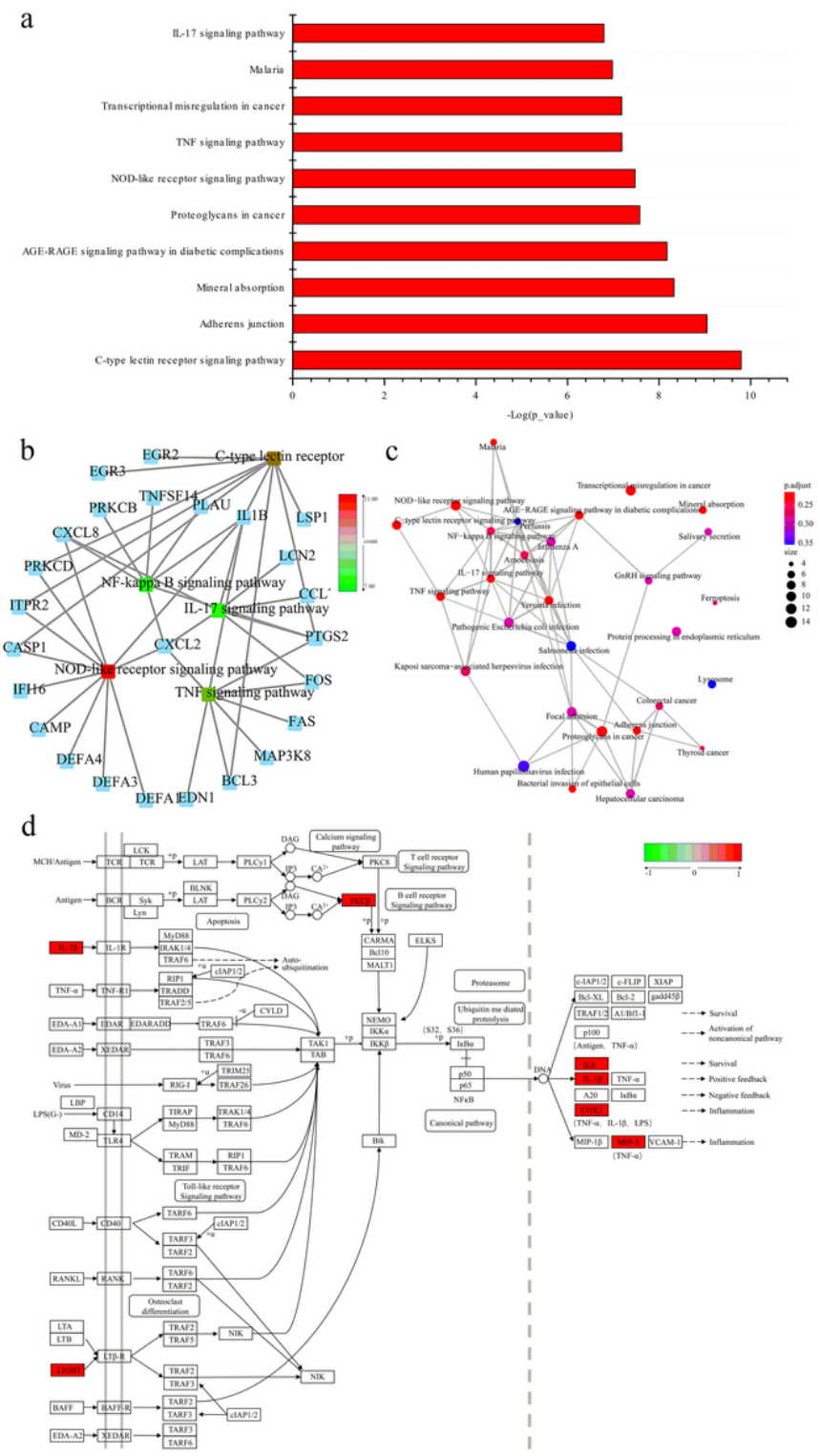

Figure 6

(a) Top 10 KEGG analysis of significantly correlated mRNAs targeted by lncRNAs. (b) Enriched gene category and top five immune pathway co-expression network. (c) Co-expression network of KEGG pathways with common genes. (d) Enriched genes of the NF-kβ signaling pathway.

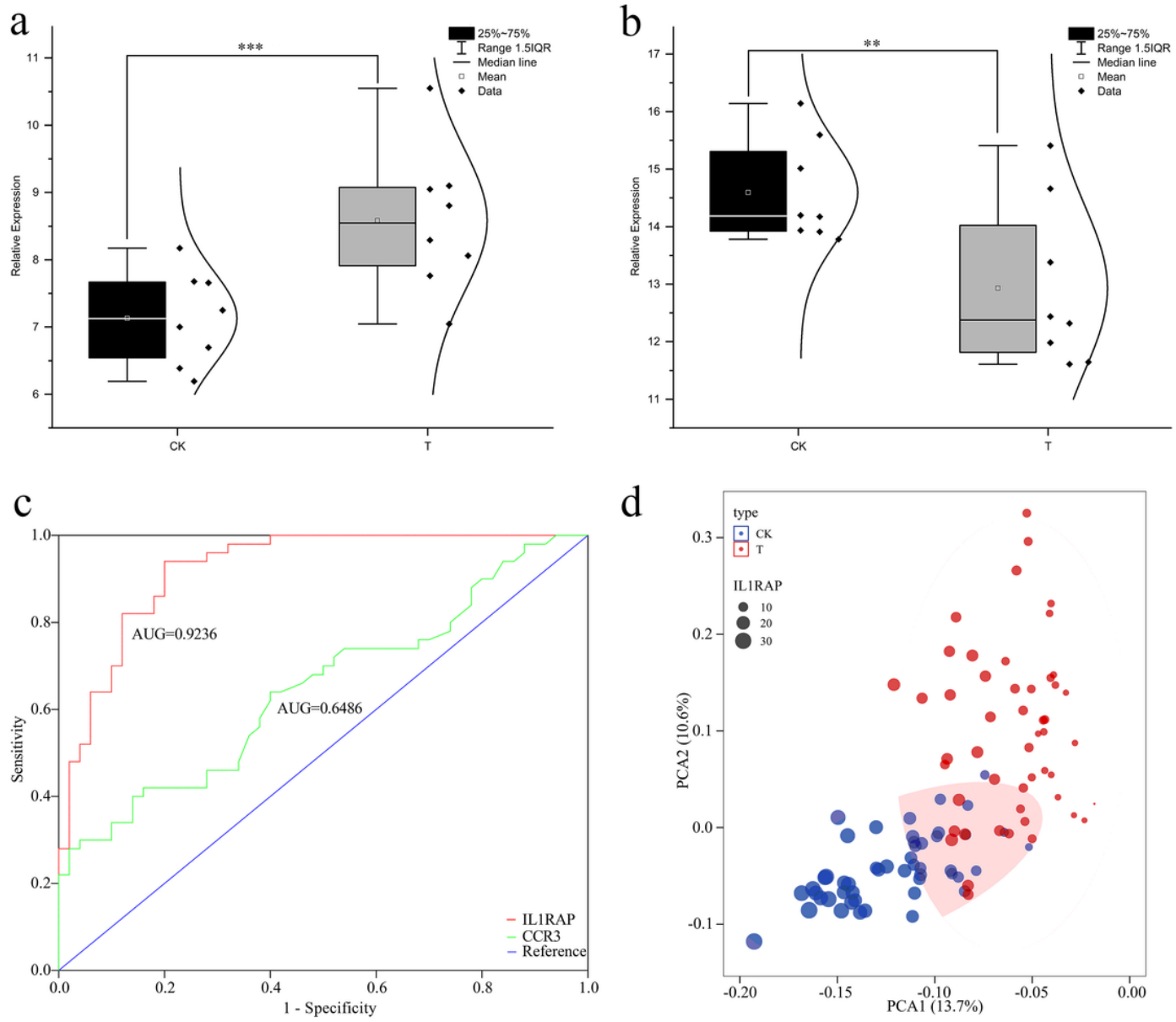

Figure 7

(a) IL1RAP-TCONS_00138311 expression level △Ct(x ± S). (b) CCR3-TCONS_00134168 expression level △Ct(x ± S). (c) ROC curves predicted by lncRNA in SCZ. (d) IL1RAP-TCONS_00138311 labeled in sample principal component analysis.

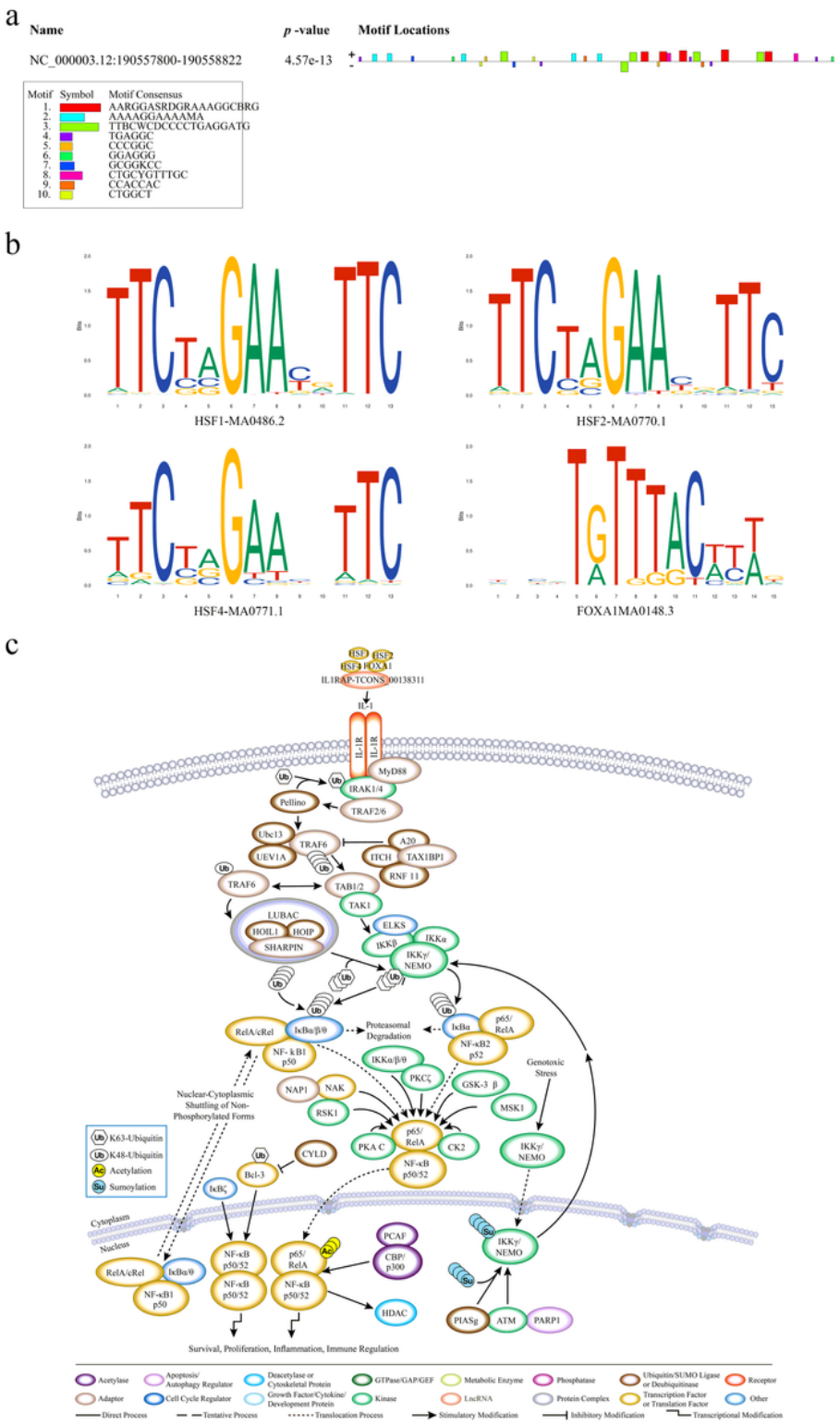

Figure 8

(a) Predicted motif_locations of IL1RAP-TCONS_00138311. (b) The sequence logo of the four most correlated transcription factors. The height of the base represents the frequency of occurrence, and the information content of the base stack represents the degree of conservation of the site. (c) IL1RAP-TCONS_00138311/IL1B/ NF-kβ action mode diagram.