

# Integrative analysis of potential biomarkers and immune cell infiltration in Parkinson's disease

Xiaopeng Chen<sup>1</sup>, Wei Cao<sup>1</sup>, Yuansu Zhuang, Siyuan Chen, Xuezhong Li<sup>\*</sup>

Department of Neurology, The Affiliated People's Hospital of Jiangsu University, Zhenjiang 221000, Jiangsu, China

## ARTICLE INFO

### Keywords:

Differentially expressed genes  
Biomarkers  
Gene expression omnibus  
Immune infiltration  
Parkinson's disease

## ABSTRACT

**Background:** Parkinson's disease (PD) is a common neurodegenerative disease in the elderly population. However, there are no reliable diagnostic biomarkers for PD, and the pathogenesis of PD still needs further study. The aim of the current study was to identify potential biomarkers and explore the pathogenesis of PD.

**Methods:** We conducted an integrative analysis of messenger RNA (mRNA), microRNA (miRNA), and long noncoding RNA (lncRNA) expression profiles of PD using data from the Gene Expression Omnibus (GEO). The GSE110720, GSE110719 and GSE133347 data sets were selected and analysed. Gene ontology (GO) enrichment and gene set variation analysis (GSVA) were performed for annotation, visualization, and integrated discovery. Protein–protein interaction (PPI) and competing endogenous RNA (ceRNA) networks were constructed, and hub genes were identified. Meanwhile, the immune infiltration analysis of hub genes was analysed. Moreover, receiver operating characteristic (ROC) curves were generated to verify the diagnostic value of the differentially expressed genes (DEGs). Finally, the genes with high area under the curve (AUC) values were verified by human samples.

**Results:** We identified 464 DEGs closely related to PD, including 154 mRNAs, 134 miRNAs, and 176 lncRNAs. The GO analyses indicated that changes in PD were mainly enriched in receptor ligand activity and cytokine receptor binding. The KEGG enrichment analysis showed that these DEGs were significantly involved in cytokine-cytokine receptor interactions, signalling pathways regulating the pluripotency of stem cells and Th17 cell differentiation. GSVA suggested that growth factor binding, IL2-stat5 signalling, and IL6-jak-stat3 signalling were crucial in the development of PD. A total of five hub genes (NPBWR2, CXCL10, CXCL5, S1PR5, and GALR1) were selected via the PPI network. A ceRNA network of the CXCL5, CXCL10 and S1PR5 genes was constructed, and target genes of the three genes were screened. The immune infiltration analysis showed that there were significant differences in a variety of immune cells between the hub genes. The expression of DEGs was validated in clinical human samples by quantitative reverse transcription polymerase chain reaction (qRT-PCR). The expression levels of hsa-miR6895–5p, hsa-miR6791–5p, hsa-miR518f–5p, hsa-miR455–3p and TEKT4P2 were decreased, while the levels of TPTE2P6 were increased in human samples. These findings are consistent with the bioinformatics analysis results.

**Conclusion:** We found that the immune inflammatory response and immune cell regulation were involved in the pathogenesis of PD. Five hub genes involved in the immune infiltration biological processes of PD based on bioinformatics. We verified the DEGs with significant differences by qRT-PCR. These findings might provide new insight into the pathogenesis of PD and the development of diagnostic and therapeutic strategies for PD.

## 1. Introduction

Parkinson's disease (PD) is a common neurodegenerative disease in the elderly population, of which common clinical manifestations are symptoms related to motion, such as static tremor, bradykinesia,

myotonia and postural balance disorder, as well as symptoms not related to motion, such as sleep disorders, loss of smell, constipation, anxiety, depression and cognitive decline. However, PD is known as a multifactorial disease due to its complex aetiology. Previous studies have shown that mitochondrial dysfunction, inflammation, oxidative stress and

<sup>\*</sup> Corresponding author.

E-mail address: [jsdxlxz@163.com](mailto:jsdxlxz@163.com) (X. Li).

<sup>1</sup> Xiaopeng Chen and Wei Cao contributed equally to this work.

apoptosis are closely related to the development of PD (Huang et al., 2019). Multiple genes have been shown to have mutations that may cause familial PD (Kalia and Lang, 2015). More than 20 gene mutation sites have been confirmed to be associated with the pathogenesis of PD (Bandres-Ciga et al., 2020). Although various efforts have been made to identify specific biomarkers of PD, there is still a lack of effective diagnostic markers (Ren et al., 2015), and the pathogenesis of PD still needs further study.

In recent years, high-throughput sequencing has been used to identify gene expression signatures, biological processes, and promising targets for PD. For example, whole genome sequencing was first used to identify all long noncoding RNAs (lncRNAs) in PD patients and found that at least 6000 kinds of lncRNAs were abnormally expressed in 2014 (Soreq et al., 2014). It is an efficient way to analyse biological data by using bioinformatics research methods, putting forward genes or gene sets related to the occurrence and development of diseases, and then carrying out experimental verification.

In this study, we used three data sets from the Gene Expression Omnibus (GEO) database for the analysis of significant differentially expressed genes (DEGs) between PD samples and normal controls and selected hub genes. A competing endogenous RNA (ceRNA) network of hub genes was constructed, and immune cell infiltration analysis was performed to explore the pathogenesis of PD. The expression of DEGs in clinical human samples was validated by quantitative reverse transcription polymerase chain reaction (qRT-PCR).

## 2. Materials and methods

### 2.1. Data acquisition and processing

We downloaded PD messenger RNA (mRNA) expression profile data (GSE110720) (PD samples=24, Normal samples=22), micro RNA (miRNA) expression profile data (GSE110719) (PD samples=9, Normal samples=6), lncRNA expression profile data (GSE133347) (PD samples=5, Normal samples=5) from NCBI GEO (<https://www.ncbi.nlm.nih.gov>) (Barrett et al., 2007). The main content of the data set is shown in Table 1 (details are provided in supplement 1). Hallmarker, C6 (genes whose gene expression changes after treatment under known conditions), Gene Ontology (GO), including biological process (BP), cellular component (CC), molecular function (MF), and Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets, were downloaded from the MSigDB database (<https://www.gsea-msigdb.org/gsea/msigdb>) (Liberzon et al., 2015). The overview of the workflow is shown in Fig. 1.

### 2.2. DEGs analysis

Since the mRNA and miRNA expression profiles of PD were count data, the edge R package was used to analyse the difference between the two data sets (the threshold was set to  $\log_2|FC| \geq 1$ ,  $p < 0.05$ ). To directly show DEGs between PD samples and normal samples, volcano maps and heatmaps were drawn by using the ggplot2 package and pheatmap package, respectively.

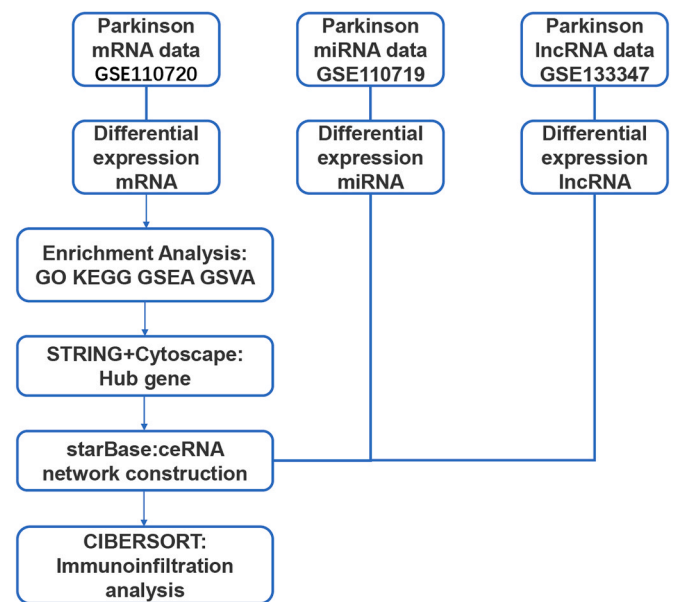


Fig. 1. The overview of the workflow.

### 2.3. Enrichment analysis

GO analysis is a common method of enrichment analysis that includes BP, MF and CC (Gene Ontology, 2015). KEGG (Kanehisa et al., 2017) is a utility database resource for understanding advanced functions and biological systems (such as cells, organisms and ecosystems). To understand which biological processes these DEGs were involved in, enrichment analysis of DEGs obtained from the mRNA data set was carried out.  $P$  value  $< 0.05$  and FDR  $< 0.25$  were used as screening thresholds.

### 2.4. Gene set variation analysis (GSVA) analysis

The GSVA package (Hanzelmann et al., 2013) was used to analyse the expression profiles of GSE110720. The reference gene sets were hallmark, GO-BP, GO-CC, GO-MF, KEGG, C6: oncogenic signatures.

### 2.5. Construction of the protein-protein interaction (PPI) network

To obtain the interaction relationship between DEGs, the DEGs obtained from GSE110720 were processed using the STRING database (<http://string-db.org/>). The interaction relationship was visualized by Cytoscape (Shannon et al., 2003) software, and the hub genes were screened by molecular complex detection.

(MCODE) (Osterman et al., 2020) and cytoHubba plug-in (Chin et al., 2014). We used a raincloud map (ggplot2 package) to show the importance of hub genes in the GO pathway.

Table 1  
Basic information of data sets.

Reference	Year	GEO ID	Platform ID	Sample type	Experiment type	PD/Control		
						Number	Age ( $\bar{x} \pm s$ )	Male
Schulze et al.	2018	GSE110720	GPL15433	Cingulate gyrus tissue	mRNA	8/8	72.5 $\pm$ 12.6/67.0 $\pm$ 15.8	4/4
				Human dermal fibroblasts	mRNA	9/6	55.8 $\pm$ 11.9/57.7 $\pm$ 12.7	8/3
				Differentiated neurons	mRNA	7/8	58.2 $\pm$ 14.3/56.0 $\pm$ 13.4	5/3
				Human dermal fibroblasts	miRNA	9/6	58.2 $\pm$ 14.3/56.0 $\pm$ 13.4	8/3
Xue Z et al.	2019	GSE133347	GPL21047	circulating leukocytes	lncRNA	5/5	61.80 $\pm$ 3.271/59.2 $\pm$ 3.3	None

## 2.6. Construction of hub gene ceRNA network

Based on starBase (<http://starbase.sysu.edu.cn/>) (Li et al., 2014), the target miRNAs of hub genes were predicted, and the predicted miRNAs were intersected with the differential miRNAs to obtain the mRNA-targeted miRNAs. In the same way, the target lncRNAs of the predicted miRNAs were also predicted and then intersected with the differential lncRNAs. All the intersection results were displayed by using the Wayne diagram, and the ceRNA network of the hub genes was constructed by Cytoscape software.

## 2.7. Correlation analysis of immune infiltration

CIBERSORT (Chen et al., 2018) immune infiltration analysis was performed on differential mRNAs of GSE110720 to see the difference in immune cells infiltration between PD and normal tissues and visualized by bar graph, correlation graph, thermograph and violin graph. Correlation analysis between five hub genes and immune infiltration cells in the mRNA data set (GSE110720).

Taking the median gene expression as the boundary, the five hub genes were divided into high and low groups, and a box diagram was used to visualize whether there was any difference in the expression of immune infiltrating cells between the high and low groups.

## 2.8. Receiver operating characteristic (ROC) analysis of DEGs

Differential mRNAs, miRNAs and lncRNAs were sorted according to the degree of difference and analysed by ROC analysis. For mRNAs, only mRNAs with area under the curve (AUC) values not less than 0.7; for miRNAs, only miRNAs with AUC values not less than 0.9; and for lncRNAs, only lncRNAs with AUC values not less than 0.95 were reserved. The first four were screened from the three data sets.

## 2.9. Clinical validation of candidate genes by qRT-PCR analysis

To verify these DEGs, we selected four miRNAs (hsa-miR6895-5p, hsa-miR6791-5p, hsa-miR518f-5p, and hsa-miR455-3p) and two lncRNAs (TEKT4P2 and TPTE2P6) for qRT-PCR validation. The study collected 55 blood samples from patients with PD (n = 30) and normal controls (n = 25). The main content of the samples is shown in Table 2 (details are provided in supplement 2). The research was approved by the Ethical Committee of the Affiliated People's Hospital of Jiangsu University (Ethics number: SH2018026) and conducted in accordance with the Declaration of Helsinki (as revised in 2013). All participants in the present study completed and submitted informed consent.

Total RNA was extracted from peripheral blood mononuclear cells (PBMCs) using TRIzol® reagent (Thermo Fisher Scientific, Inc.). RNA quality was then determined with a spectrometer (Nano Drop 2000, Thermo Scientific, Wilmington, DE, USA), and the assessment of purity relied on the A260/A280 ratio. Then, reverse transcription was performed with a reverse transcription kit (Takara), and primers were designed and purchased from Sangon Biotech. Quantitative PCR was

**Table 2**

Clinical and demographic characteristics of the PD patients.

	PD	Control
Number	30	25
Sex(Male)	12	13
Age(±s)	63.6 ± 5.8	66.9 ± 6.6

CRediT authorship contribution statement

Xuezhong Li conceived and designed the experiments; Yuansu Zhuang and Siyuan Chen collected and analysed the patients information; Siyuan Chen and Xiaopeng Chen analysed and interpreted statistical data. Wei Cao and Xiaopeng Chen carried out bioinformatics analysis, PCR analysis and drafted the manuscript.

performed on an ABI 7500 Real-time PCR with a fluorescent quantitative PCR kit (Takara). The study analysed the relative gene expression by using the  $2^{-\Delta\Delta Ct}$  method with human GAPDH serving as an endogenous control for the expression of genes. The following primers were used: hsa-miR6895-5p-forward 5'-CAGGGCCAGGCACAGAGTAAG-3', hsa-miR6791-5p-forward 5'-TATATATATATCCCCTGGGGCTGGGC-3', hsa-miR518f-5p-forward 5'-CGCTCTAGAGGAAGCACTTTCTC-3', hsa-miR455-3p-forward 5'-CCGCAGTCCATGGGCATATACAC-3', TEKT4P2-forward 5'-CCCTCT.

CCAATAAACACTCCAATCTCC-3' and TEKT4P2- reverse 5'-TTTAAGGCGTGCT.

TGCATGTTGTG-3', TPTE2P6- forward 5'-AGCATCAGGAACAAGA-GAAG-3' and TPTE2P6- reverse 5'-TCAGCAACACATCAGTAACA -3'. The four miRNA primers were designed by tailing methods. All experiments were performed in triplicate.

## 2.10. Statistical analysis method

All analyses were carried out by R software (version 3.6.2), and *P* values less than 0.05 were considered significant differences. The obtained qRT-PCR data are presented as the mean ± standard error (SEM). Student's *t*-test or one-way ANOVA was used to compare the differences between groups. The statistical analysis was conducted with SPSS version 25.0. A value of *P* < 0.05 indicated statistically significant results.

## 3. Results

### 3.1. Identification of DEGs

Compared with normal samples, a total of 154 DEGs (81 upregulated and 73 downregulated genes) in GSE110720, 134 DEGs (27 upregulated and 107 downregulated genes) in GSE110719, and 176 DEGs (112 upregulated and 64 downregulated genes) in GSE133347 were identified in PD samples (Fig. 2).

### 3.2. Enrichment analysis of mRNA data set

The 154 differentially expressed mRNAs were enriched by GO (Fig. 3) and KEGG (Fig. 4) analyses. GO enrichment analysis showed that these genes were significantly enriched in receptor ligand activity, cell fate commitment, cytokine receptor binding, and central nervous system neuron differentiation. KEGG enrichment analysis showed that these genes were significantly enriched in cytokine-cytokine receptor interactions, Hippo signalling pathways, signalling pathways regulating pluripotency of stem cells and Th17 cell differentiation.→.

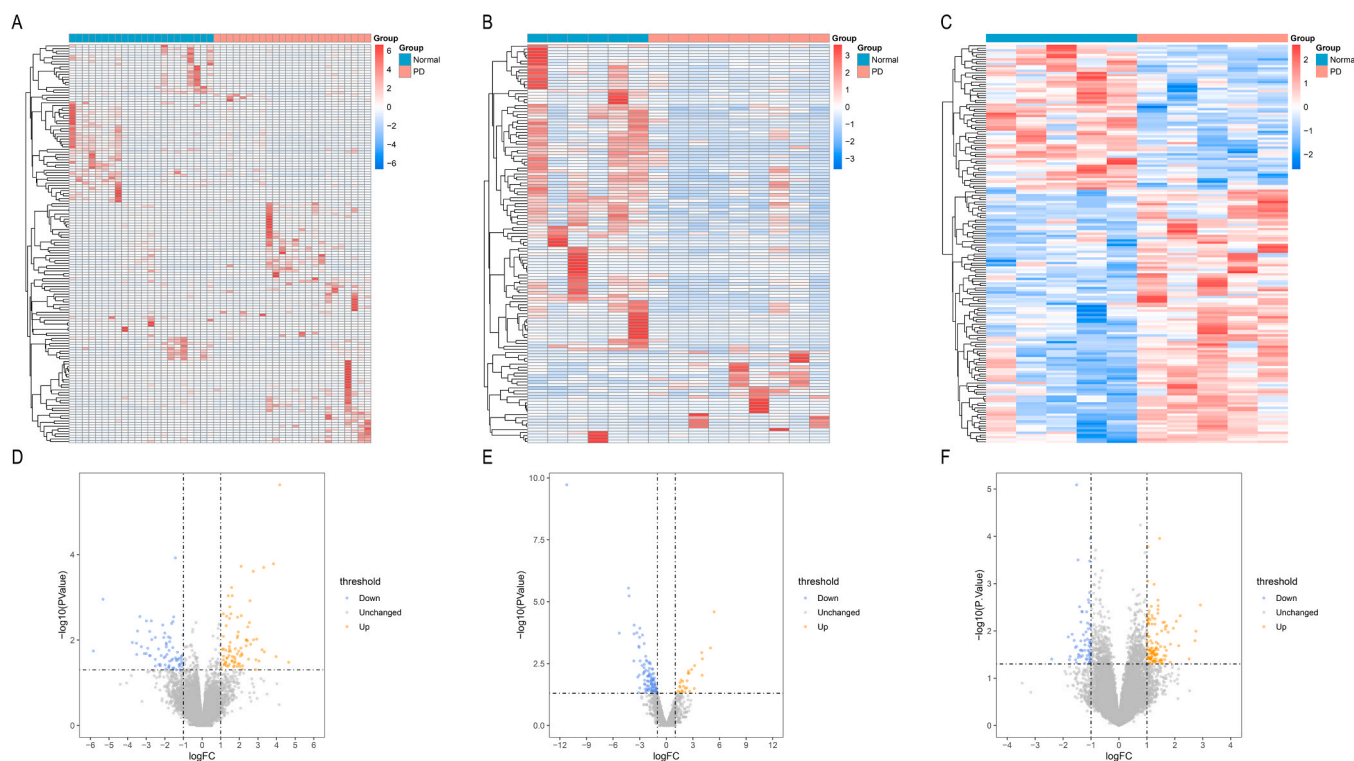
### 3.3. GSVA analysis of mRNA data set

GSVA analysis was performed on the mRNA expression profile GSE110720 data set.

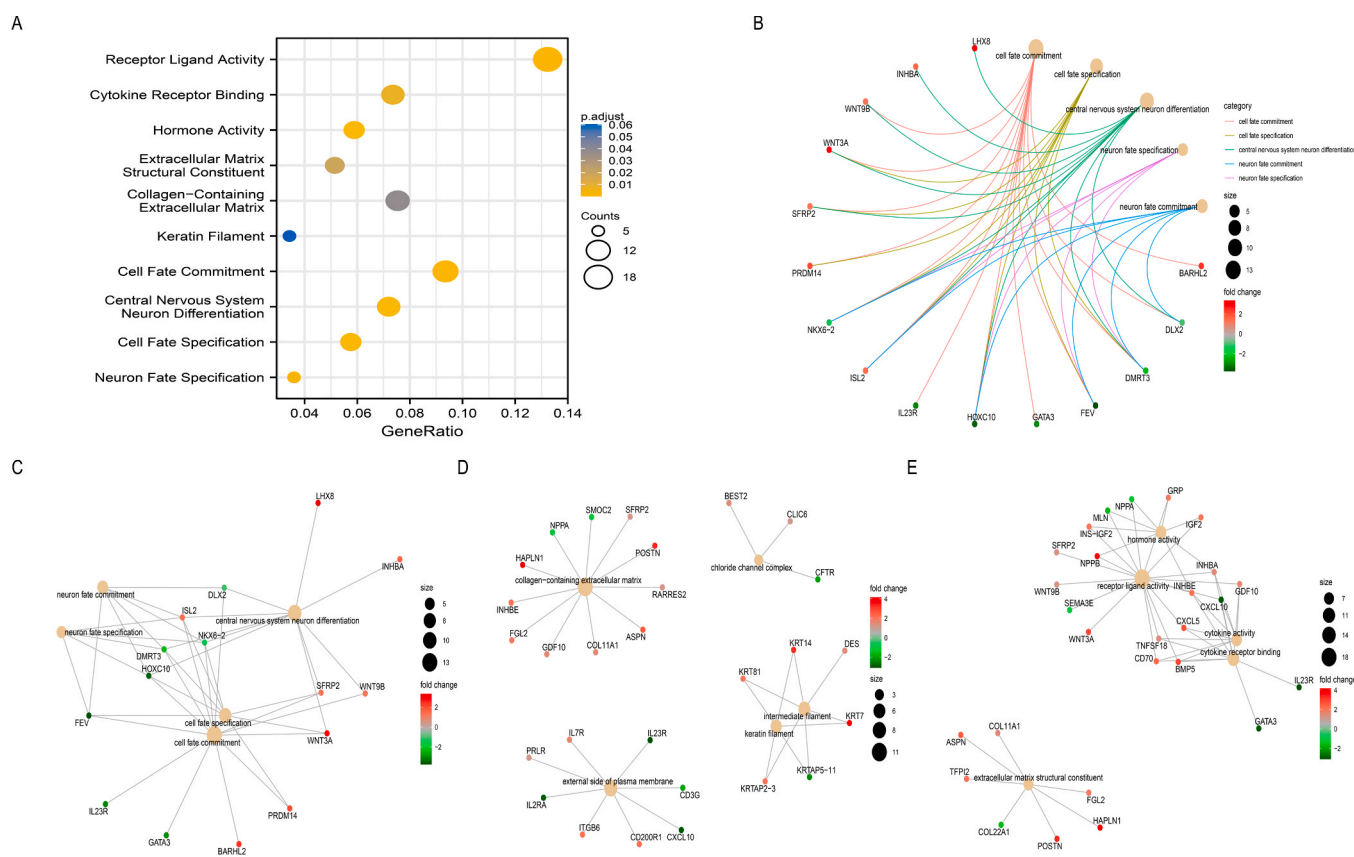
Taking different gene sets as reference sets, the pathways of these reference sets significantly enriched between PD tissues and normal tissues in the GSE110811 data set were obtained (Fig. 5). The results showed that these genes were significantly enriched in growth factor binding, transferase activity transferring nitrogenous groups, p53 binding, central nervous system development, protein maturation, primary bile acid biosynthesis, leukocyte transendothelial migration, hedgehog signalling pathway, IL6-jak-stat3 signalling, epithelial mesenchymal transition, and IL2-stat5 signalling.

### 3.4. Construction of PPI network of mRNA data set

For these 154 differential mRNAs, we constructed an interaction network (154 differential genes, 74 nodes and 119 edges after removing outliers) (Fig. 6A). The network showed that the larger the degree of



**Fig. 2.** Identification of DEGs. (A) Difference analysis of GSE110720 visualized by heat maps. (B) Difference analysis of GSE110719 visualized by heat maps. (C) Difference analysis of GSE133347 visualized by heat maps. (D) Upregulated and downregulated genes of GSE110720 visualized by volcano maps. (E) Upregulated and downregulated genes of GSE110719 visualized by volcano maps. (F) Upregulated and downregulated genes of GSE133347 visualized by volcano maps.



**Fig. 3.** GO enrichment analysis. (A) GO enrichment analysis shown by bubble diagram. (B) GO enrichment analysis shown by circle diagram. (C) GO-BP enrichment analysis. (D) GO-CC enrichment analysis. (E) GO-MF enrichment analysis.



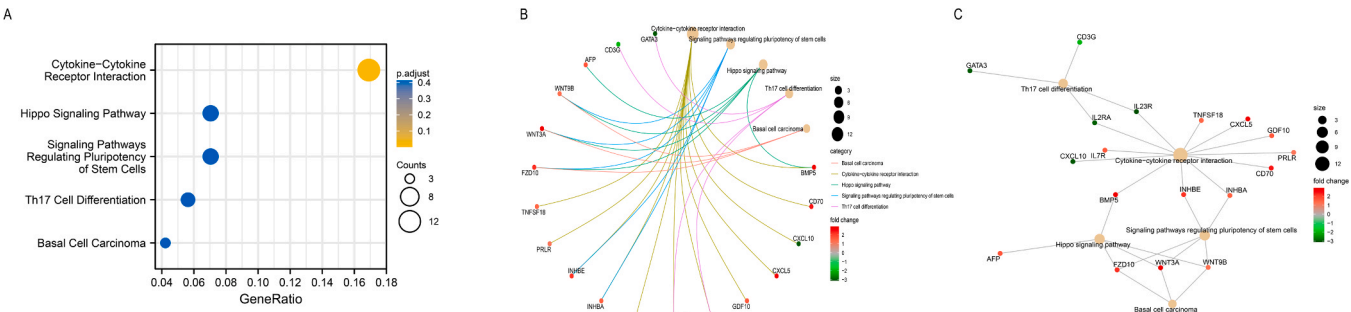


Fig. 4. KEGG analysis. (A) KEGG analysis bubble diagram. (B) KEGG analysis circle diagram. (C) The network diagram of KEGG analysis.

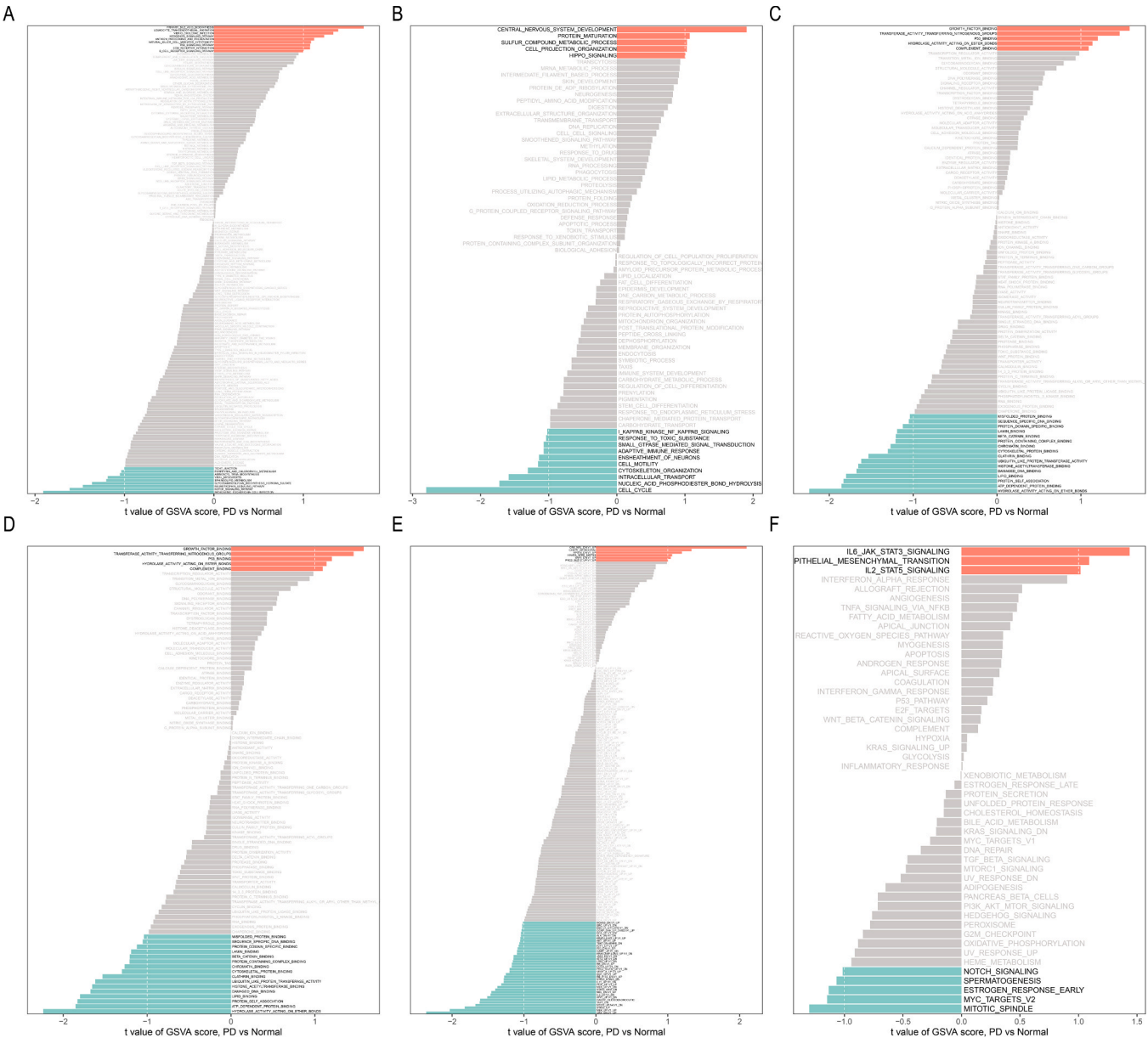
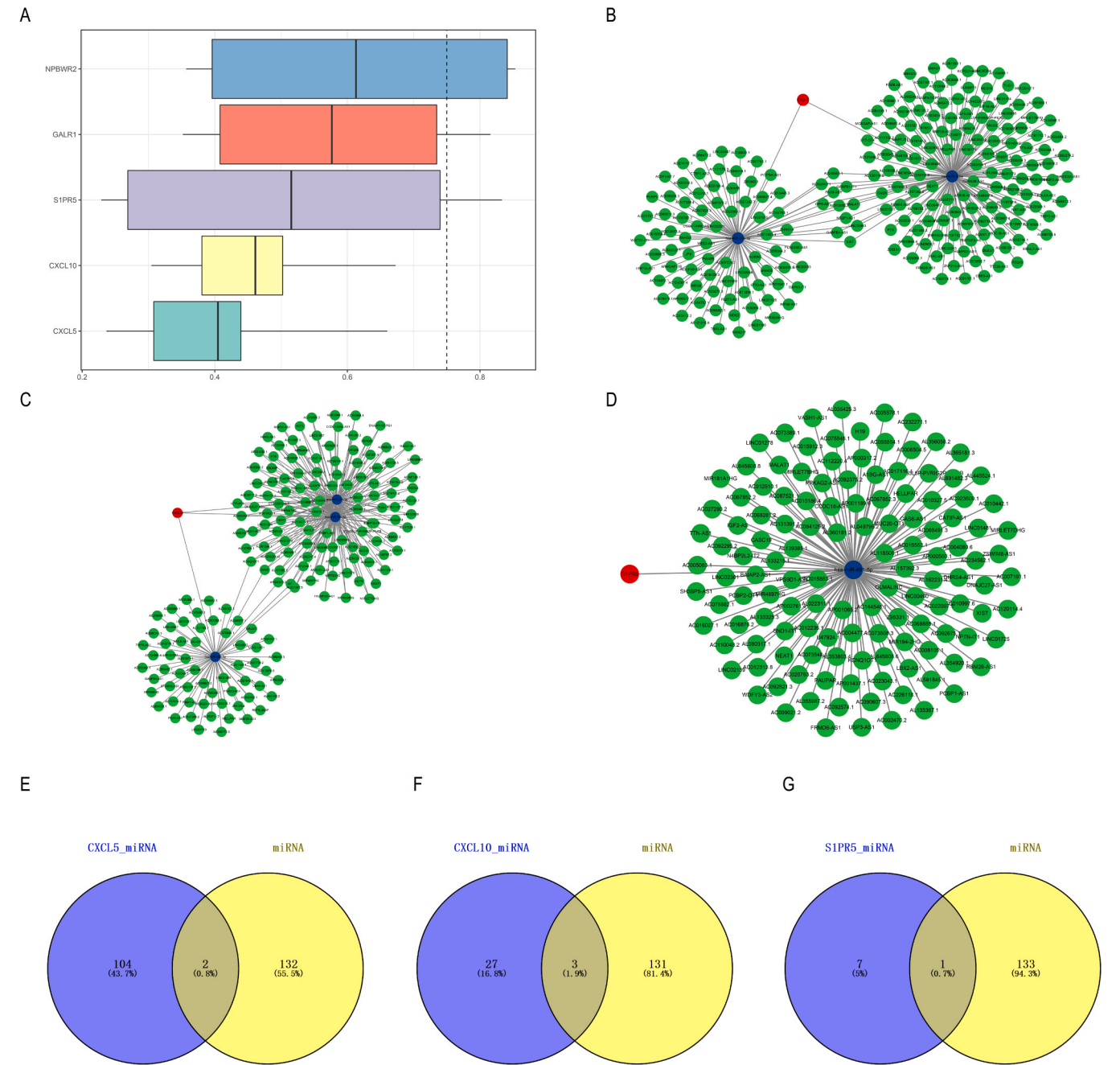
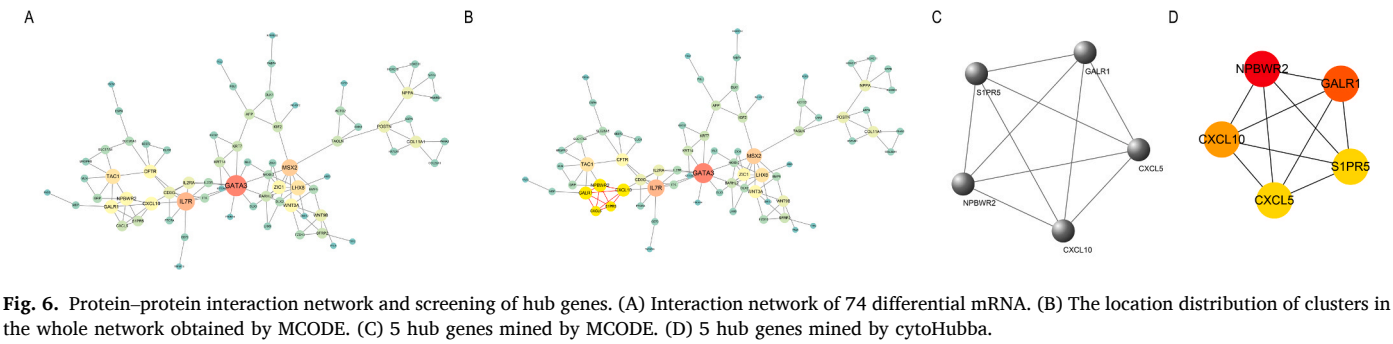


Fig. 5. Results of different reference gene sets of GSVA. (A) KEGG gene sets. (B) GO-BP gene sets. (C) GO-CC gene sets. (D) GO-MF gene sets. (E) Oncogenic signatures gene sets. (F) Hallmarker gene sets.

nodes, the redder the colour and the larger the points. Using MCODE, we obtained the closest cluster of PPIs (5 nodes, 10 edges, MCODE score  $\geq 4$ ) (Fig. 6B). The yellow part was the cluster of MCODE. The nodes in this

cluster were highly connected (Fig. 6C). Meanwhile, using the cytoHubba mining module, the first five genes were taken as the hub genes of the whole PPI (5 nodes, 10 edges) (Fig. 6D). The hub genes mined by



MCODE and cytoHubba were identical.

### 3.5. Construction of hub genes ceRNA network

To further select the more important genes, we evaluated the importance of five hub genes based on the three semantics of BP, CC and MF of GO. As shown in Fig. 7A, NPBWR2 was the most important gene, while CXCL5 was the least important. Based on the starBase database, the target genes of five hub genes were predicted, among which only the CXCL5, CXCL10 and S1PR5 genes could successfully predict the targeted miRNAs. After the intersection between these targeted miRNAs and differential miRNAs of GSE110719, two, three, and one predicted differential miRNA were obtained (Fig. 7E–G). Interestingly, the predicted lncRNAs and differential lncRNAs of GSE133347 were different when we took the intersection. Based on the above, we constructed a ceRNA network of mRNA - predicted-differential - miRNAs predicted lncRNA for CXCL5, CXCL10 and S1PR5 genes (Fig. 7B–D).

### 3.6. Immune infiltration in mRNA data set

The mRNA data set (GSE110720) was analysed for immune infiltration. The composition of immune cells in each sample was analysed first and displayed by a bar graph (Fig. 8A). Then, the correlation of the expression of immune cells in the data set was analysed. The results showed that M2 macrophages, monocytes, resting NK cells, activated mast cells, and naive T cells had a highly positive correlation. However, memory CD4 T cells, M2 macrophages, and regulatory T cells had obviously negative correlations (Fig. 8B). Finally, the difference in the expression of immune cells between the disease and normal samples was analysed and displayed as a heat map and violin map. The results showed that B cell memory was significantly decreased in PD samples (Fig. 8C–D).

### 3.7. The relationship between hub genes and immune infiltration

The correlation analysis between five hub genes and immune infiltrating cells in the mRNA (GSE110720) data set showed that there were significant differences in a variety of immune cells between the high and low expression groups of each gene (Fig. 9), especially in resting memory CD4 T cells, naive CD4 T cells, monocytes, resting NK cells, neutrophils, and naive B cells.

### 3.8. ROC analysis of DEGs

According to the threshold value of AUC 0.7, 0.9 and 0.95, the four most significant mRNAs (SYNGR4, PRLR, POTEI, NPPA) (Fig. 10A–D), miRNAs (hsa-miR6895–5p, hsa-miR6791–5p, hsa-miR518f–5p, hsa-miR455–3p) (Fig. 10E–H) and lncRNAs (ZMYND10-AS1, TPTE2P6, TEK4P2, SNORD27) (Fig. 10I–L) of DEGs associated with PD were screened out.

### 3.9. Validating the expression of candidate genes by RT-qPCR

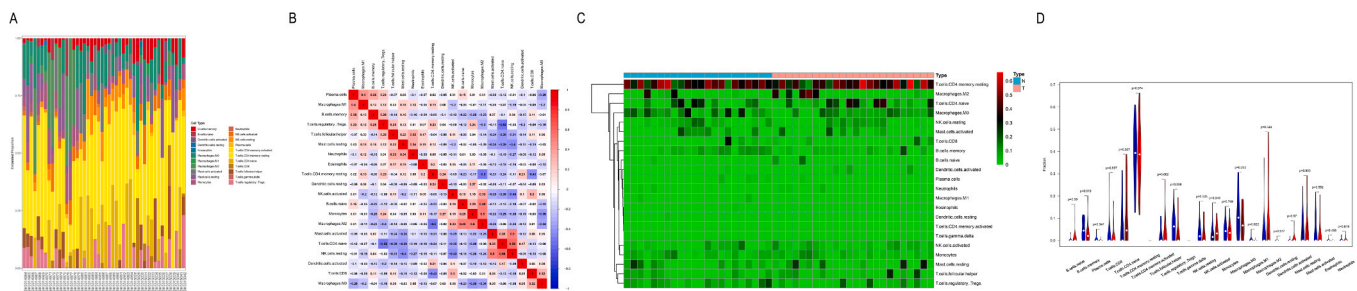
We selected four miRNAs (hsa-miR6895–5p, hsa-miR6791–5p, hsa-miR518f–5p, hsa-miR455–3p) and two lncRNAs (TEK4P2, TPTE2P6) as candidate genes for validation by qRT-PCR. The  $2^{-\Delta\Delta Ct}$  function was used to calculate differential gene expression. The  $\log_2(2^{-\Delta\Delta Ct})$  value was used to determine the relative expression of DEGs. The results showed that the expression of five genes (hsa-miR6895–5p, hsa-miR6791–5p, hsa-miR518f–5p, hsa-miR455–3p, TEK4P2) was down-regulated, while only TPTE2P6 expression was upregulated (Fig. 11).

## 4. Discussion

With the increase in the PD incidence rate, understanding the pathological and molecular mechanisms of PD is critical for clinical diagnosis and treatment. Efficient high-throughput sequencing and bioinformatics analyses have contributed to our understanding of the molecular mechanisms of disease occurrence and development, which is necessary to explore genetic alterations and identify potential diagnostic biomarkers.

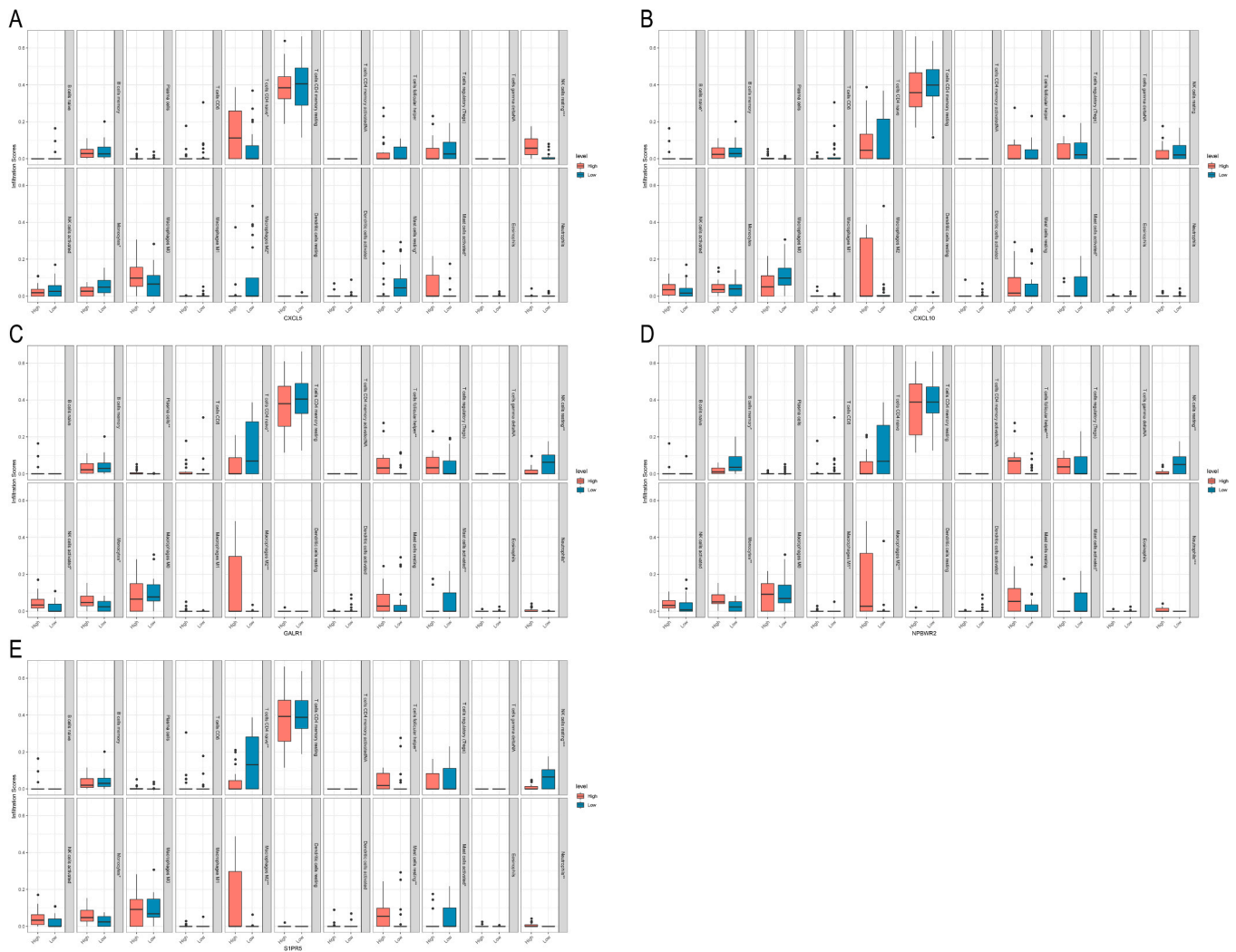
Therefore, in this study, we analysed the high-throughput sequencing and microarray data of three groups of PD patients (38 cases in total) and normal controls (31 cases in total) by bioinformatics methods, including mRNA, miRNA and lncRNA data sets. First, we identified 154 DEGs (mRNAs) in the GSE110720 data set, including 81 upregulated genes and 73 downregulated genes. GO analyses revealed that the 154 differentially expressed mRNAs were mostly enriched in receptor ligand activity and cytokine receptor binding. KEGG enrichment analysis showed that these DEGs were significantly involved in cytokine-cytokine receptor interactions, signalling pathways regulating the pluripotency of stem cells and Th17 cell differentiation. Notably, these genes were mainly enriched in the immune inflammatory response and immune cell regulation. A substantial body of evidence suggests that the immune system is involved in PD pathogenesis (Tan et al., 2020). PD patients have elevated proinflammatory cytokines, such as interferon (IFN) $\gamma$ , interleukin (IL)–1 $\beta$ , IL-2, IL-6 and tumour necrosis factor  $\alpha$  (TNF $\alpha$ ), in the blood and cerebrospinal fluid (CSF), which contribute to neuronal death (Karpenko et al., 2018). Immune cell dysregulation and perturbations in immune subsets in PD patients through immunophenotyping reveal intrinsic changes, leading to an inflammatory predisposition. The central involvement of immune mediators arises from the breakdown of the blood brain barrier (BBB), which results in the presence of complement proteins, inflammasome activation, infiltration of CD4<sup>+</sup> and CD8<sup>+</sup> T lymphocytes and autoantibodies against neuronal antigens in the substantia nigra (SN) of PD patients (Chen et al., 2018). However, the precise molecular mechanism leading to dopaminergic neuronal death through immune pathways remains unclear.

To investigate the biological functions of the DEGs associated with PD, GSVA was performed using the GSVA package, and the common activated and suppressed signalling pathways were selected from GEO110720. The results showed that the genes were significantly



**Fig. 8.** Correlation analysis of key mRNA with immune infiltration. (A) The bar chart of immune cell content of each sample was analysed in GSE110720 data set. (B) Correlation diagram of various immune cell contents. (C) Cluster heat map of immunocyte content in normal and PD samples. (D) Violin plot of immunocyte content in normal and PD samples (the blue was the normal samples and the red was the PD samples).





**Fig. 9.** Correlation analysis of 5 hub genes with immune infiltration. (A) The correlation between CXCL5 and immune infiltration. (B) The correlation between CXCL10 and immune infiltration. (C) The correlation between GALR1 and immune infiltration. (D) The correlation between NPBWR2 and immune infiltration. (E) The correlation between S1PR5 and immune infiltration.

enriched in growth factor binding, IL2-stat5 signalling, and IL6-jak-stat3 signalling. IL2 is a pleiotropic cytokine, and dissecting the signalling pathways that allow IL2 to control the differentiation and homeostasis of both pro- and anti-inflammatory T cells is fundamental to determining the molecular details of immune regulation. The IL2 receptor couples to JAK tyrosine kinases and activates STAT5 transcription factors; it is a key regulator of T cell metabolic programmes (Ross and Cantrell, 2018). IL6 is secreted by T cells and is a proinflammatory factor. IL6-jak-stat3 signalling is a classical inflammatory pathway in our bodies. In PD patients, Th17 cells are regulated by IL6 and signal transducer and activator of transcription 3 (STAT3) (Yeung et al., 2018). Meanwhile, IL-6 signalling cascades can interfere with growth factor (GF) receptor signalling pathways, such as nerve growth factor (NGF) receptor signalling (Prots and Winner, 2019). The above two pathways in PD may play a crucial role in regulating immunity and inflammation.

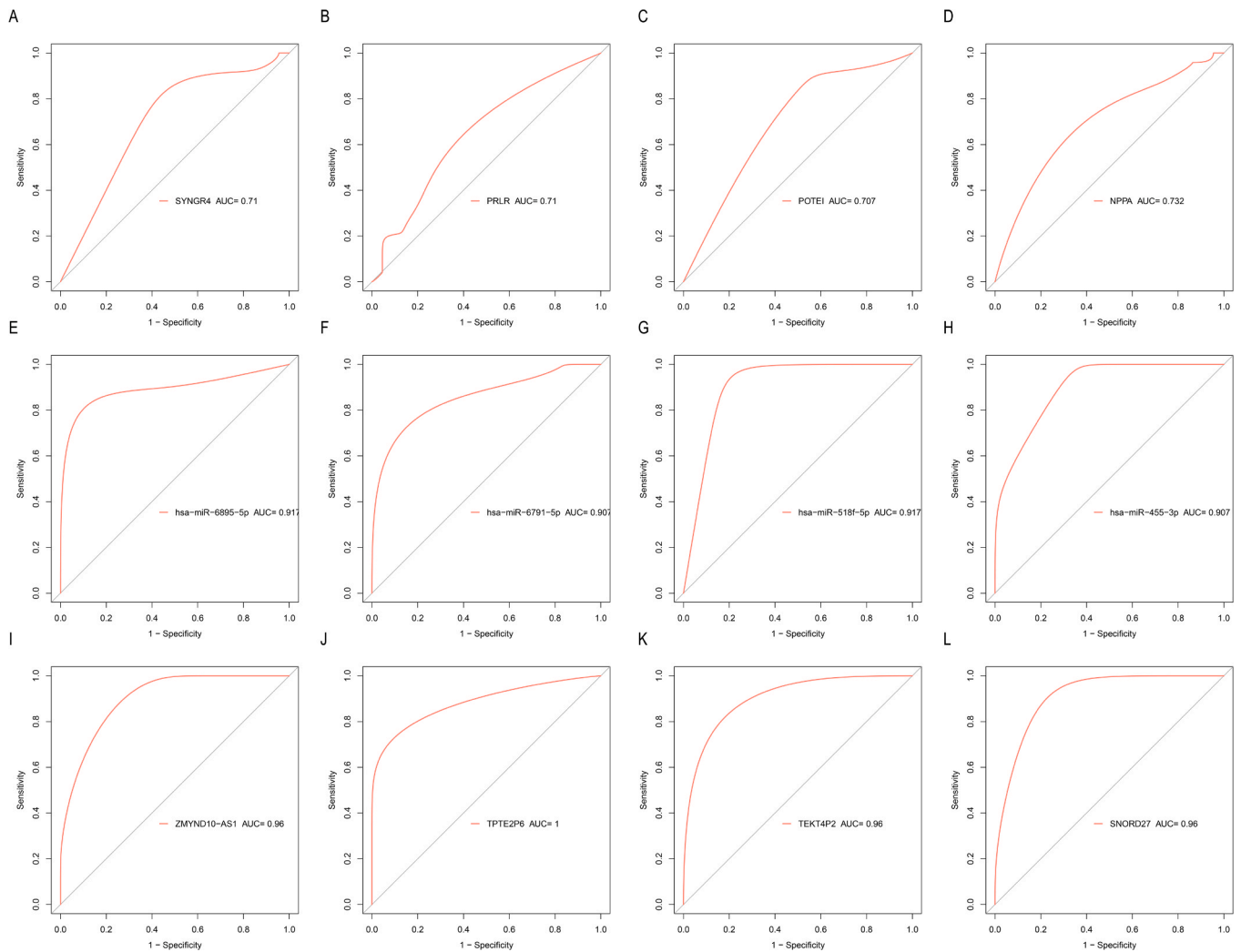
In the present study, five overlapping genes, NPBWR2, CXCL10, CXCL5, S1PR5, and GALR1 in GSE110720 were selected as the most significant hub genes using the PPI network. The NPBWR2 gene was highly expressed in several brain regions, including the hypothalamus and midbrain. Neuropeptide B (NPB) is an endogenous neuropeptide ligand for NPBWR2 and plays a variety of roles through the receptor ligand pathway. The majority of NPB neurons in the mesolimbic region possess tyrosine hydroxylase immunoreactivity, indicating that a small

subset of dopaminergic neurons coexpress NPB (Chottova Dvorakova, 2018). CXCL10 and CXCL5 belong to the chemokine family, in which CXCL10 is a T cell chemokine. CXCL10 expression correlated with the tissue infiltration of T cells. Chemokines are expressed and contribute to defence against chronic diseases such as PD by attracting populations of leukocytes to the central nervous system (CNS) to activate T and B lymphocyte expression (Koper et al., 2018). GALR1 was mainly distributed in the CNS and gastrointestinal tract, and galanin (GAL) was the ligand of GALR1. They produced effects through receptor ligand activity. GAL was suggested to have an anti-inflammatory function in some situations in the CNS (Oliveira Volpe et al., 2020).

Sphingosine 1-phosphate receptors (S1PR) are G protein-coupled, and S1PR5 is mainly expressed in the brain. Fingolimod targeted S1PRs and was the first oral therapy for patients with relapsing-remitting multiple sclerosis (MS). A study showed that fingolimod attenuated CXCL10 and CXCL5 protein release from astrocytes (O'Sullivan et al., 2018).

In this study, three hub genes were successfully constructed in the ceRNA network, including CXCL10, S1PR5 and CXCL5. Further research regarding the relationship found that six miRNAs play an essential role in PD, including hsa-miR-29c-3p, hsa-miR-29b-3p, hsa-miR-491-5p, hsa-miR-92a, hsa-miR-320a, and hsa-miR-651-5p. miR-29c-3p modulates the NOD-like receptor protein 3 (NLRP3) inflammasome to impair





**Fig. 10.** ROC analysis of key mRNA, miRNA and lncRNA. (A) ROC curve of the SYNGR4. (B) ROC curve of the PRLR. (C) ROC curve of the POTEI. (D) ROC curve of the NPPA. (E) ROC curve of the hsa-miR-6895-5p. (F) ROC curve of the hsa-miR-6791-5p. (G) ROC curve of the hsa-miR-518f-5p. (H) ROC curve of the hsa-miR-455-3p. (I) ROC curve of the ZMYND10-AS1. (J) ROC curve of the TPTE2P6. (K) ROC curve of the TEK4P2. (L) ROC curve of the SNORD27.

microglial inflammatory responses by targeting nuclear factor 5 of activated T cells (NFAT5), which represents a promising therapeutic target for PD (Wang et al., 2020). miR-491 caused a marked reduction in dopamine transporter (DAT) expression, influencing neuronal dopamine transport. Moreover, the regulation of miR-491 on this transport disappeared after DAT was silenced, suggesting that miR-491 negatively regulates dopamine transporter expression and function in neural cells (Jia et al., 2016). hsa-miR-92a was found to be the common hub between regulatory and coexpression networks through bioinformatics analysis by Chatterjee et al., suggesting its strong functional role in PD (Chatterjee et al., 2014). However, there were still no reports on the correlation between the genes of hsa-miR-320a and hsa-miR-651-5p with PD.

CIBERSORT (Newman et al., 2015) has been widely used to evaluate the relative content and dynamic regulation process of 22 kinds of immune cells. In our study, the DEGs we found were mainly enriched in the immune inflammatory response and immune cell regulation through GO and KEGG analysis. We found that memory B cells of 22 kinds of immune cells were significantly different between the PD and normal control groups. A study found that deposits of IgG are found on dopaminergic neurons in these patients, and Lewy bodies themselves are coated with IgG, which suggests that B cells might be involved in the pathogenesis of PD (Sabatino et al., 2019). When the expression levels of five hub genes were used to analyse the types and distribution of various immune cells,

they were all involved in the immune process, and the distribution of immune cells among the genes was similar. For example, among the five hub genes, there was an increase in the proportion of monocytes. The monocytes in the brain were mainly microglia, accounting for 5% of the total number of glial cells. In the presence of infection, microglia were activated and transformed into monocytes. Inflammation can upregulate the expression of CXCL10, CXCL12 and other genes in microglia and aggravate the progression of PD (Li et al., 2019). In the CXCL5, GALR1 and S1PR5 genes, there was a decrease in the proportion of naive CD4 T cells, which was consistent with Fiszer reports (Fiszer et al., 1994). Follicular helper T cells (T<sub>fh</sub>s) are a new type of T cell subset that play an important role in stimulating B cell proliferation, differentiation and immunoglobulin conversion (Mak et al., 2017). This study found that the proportion of T<sub>fh</sub>s increased in the GALR1, NPBWR2 and S1PR5 genes, indicating that T<sub>fh</sub>s play an important role in the immune inflammatory response in PD.

The four most significant miRNAs and lncRNAs associated with PD showed significant predictive accuracy by ROC curve analysis, and all AUCs were above 0.9, suggesting that these new biomarkers can be further studied in PD. We selected six genes for RT-qPCR in our own tissue samples to validate the integrated data. The results showed that these genes were significantly different between patients with PD and normal controls and may serve as novel potential targets for the diagnosis and treatment of PD.

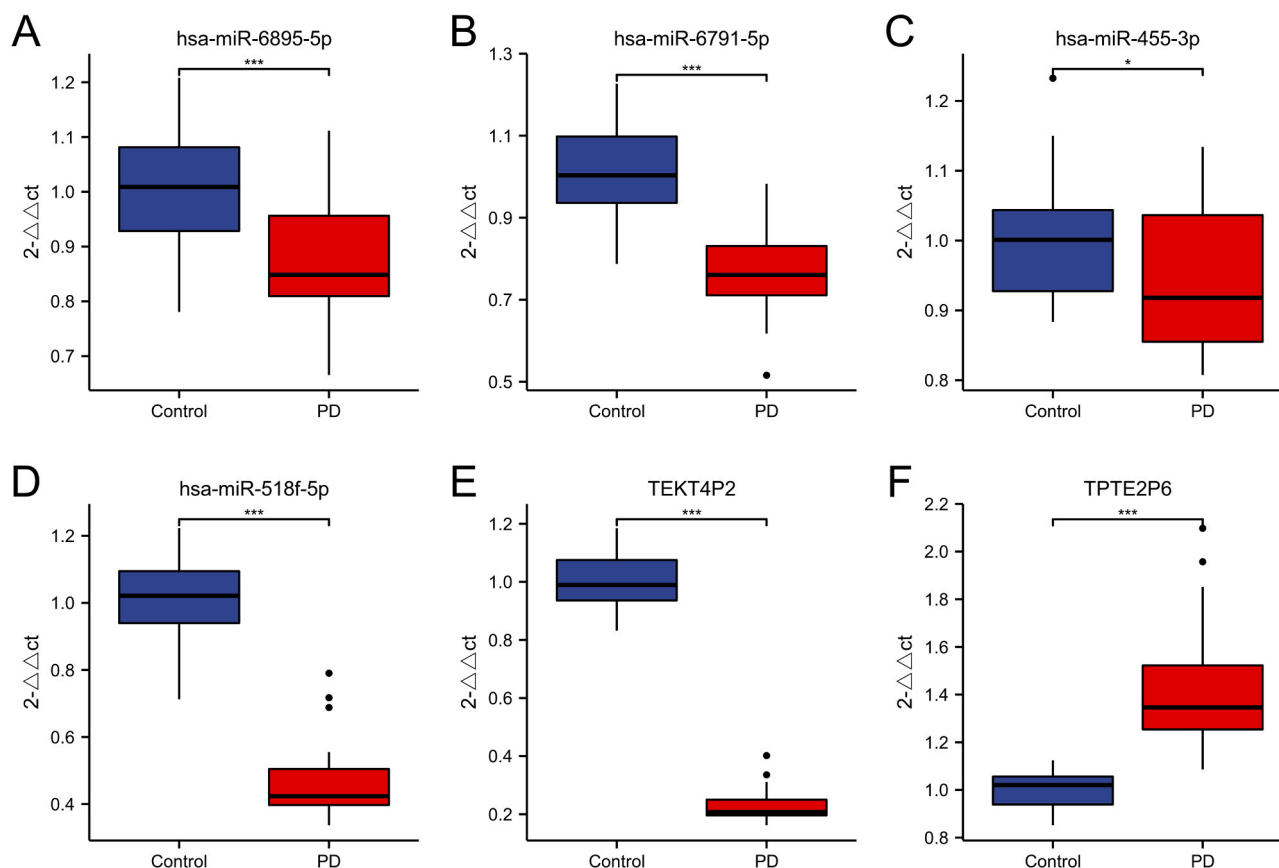


Fig. 11. Clinical validation of genes of interest by qRT-PCR (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ).

Despite the significant results obtained in the present study, there were several shortcomings. First, the three sets of data are from three single GEO databases, and a single analysis may contribute to a high false-positive rate and one-sided results. Second, our data are from public databases, and the sample size is too small to evaluate the data quality. Moreover, lncRNAs are not from high-throughput sequencing but from microarrays, and many unknown molecules are still undetected. In addition, we have only identified a few of the molecules that are most relevant to PD in small human samples, and whether they are related to the course of disease and the severity of the disease needs further verification. Meanwhile, animal and cell experiments are needed to further prove these results.

In summary, our current study provided a reliable comprehensive analysis by combining three data sets, including mRNA, miRNA, and lncRNA. In total, 154 DEGs in GSE110720, 134 DEGs in GSE110719, and 176 DEGs in GSE133347 were identified. GO, KEGG and GSEA analyses demonstrated that inflammatory responses played an important role in the pathogenesis of PD. The immune infiltration analysis of 5 PD genes, which could be potential diagnostic biomarkers through PPI network analysis, also verified the above view. The genes verified by our RT-qPCR may become new biomarkers in the future.

#### CRediT authorship contribution statement

Xuezhong Li conceived and designed the experiments; Yuansu Zhuang and Siyuan Chen collected and analysed the patients information; Siyuan Chen and Xiaopeng Chen analysed and interpreted statistical data. Wei Cao and Xiaopeng Chen carried out bioinformatics analysis, PCR analysis and drafted the manuscript.

#### Conflict of interest

The authors declare that they have no conflicts of interest.

#### Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.brainresbull.2021.09.010](https://doi.org/10.1016/j.brainresbull.2021.09.010).

#### References

- Bandres-Ciga, S., Diez-Fairen, M., Kim, J.J., Singleton, A.B., 2020. Genetics of Parkinson's disease: an introspection of its journey towards precision medicine. *Neurobiol. Dis.* 137, 104782. <https://doi.org/10.1016/j.nbd.2020.104782>.
- Barrett, T., Troup, D.B., Wilhite, S.E., Ledoux, P., Rudnev, D., Evangelista, C., Kim, I.F., Soboleva, A., Tomashevsky, M., Edgar, R., 2007. NCBI GEO: mining tens of millions of expression profiles—database and tools update (Database issue). *Nucleic Acids Res.* 35, D760–D765. <https://doi.org/10.1093/nar/gkl887>.
- Chatterjee, P., Bhattacharyya, M., Bandyopadhyay, S., Roy, D., 2014. Studying the system-level involvement of microRNAs in Parkinson's disease. *PLoS One* 9 (4), 93751. <https://doi.org/10.1371/journal.pone.0093751>.
- Chen, Z., Chen, S., Liu, J., 2018. The role of T cells in the pathogenesis of Parkinson's disease. *Prog. Neurobiol.* 169, 1–23. <https://doi.org/10.1016/j.pneurobio.2018.08.002>.
- Chen, B., Khodadoust, M.S., Liu, C.L., Newman, A.M., Alizadeh, A.A., 2018. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* 1711, 243–259. [https://doi.org/10.1007/978-1-4939-7493-1\\_12](https://doi.org/10.1007/978-1-4939-7493-1_12).
- Chin, C.H., Chen, S.H., Wu, H.H., Ho, C.W., Ko, M.T., Lin, C.Y., 2014. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8, 11. <https://doi.org/10.1186/1752-0509-8-S4-S11>.
- Chottova Dvorakova, M., 2018. Distribution and function of neuropeptides W/B signaling system. *Front. Physiol.* 9, 981. <https://doi.org/10.3389/fphys.2018.00981>.
- Fiszer, U., Mix, E., Fredrikson, S., Kostulas, V., Link, H., 1994. Parkinson's disease and immunological abnormalities: increase of HLA-DR expression on monocytes in cerebrospinal fluid and of CD45RO+ T cells in peripheral blood. *Acta Neurol. Scand.* 90 (3), 160–166. <https://doi.org/10.1111/j.1600-0404.1994.tb02699.x>.
- Gene Ontology, C., 2015. Gene ontology consortium: going forward (Database issue). *Nucleic Acids Res.* 43, D1049–D1056. <https://doi.org/10.1093/nar/gku1179>.

- Hanzelmann, S., Castelo, R., Guinney, J., 2013. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* 14, 7. <https://doi.org/10.1186/1471-2105-14-7>.
- Huang, M.L., Chiang, S., Kalinowski, D.S., Bae, D.H., Sahni, S., Richardson, D.R., 2019. The role of the antioxidant response in mitochondrial dysfunction in degenerative diseases: cross-talk between antioxidant defense, autophagy, and apoptosis. *Oxid. Med. Cell. Longev.* 2019, 6392763 <https://doi.org/10.1155/2019/6392763>.
- Jia, X., Wang, F., Han, Y., Geng, X., Li, M., Shi, Y., Lu, L., Chen, Y., 2016. miR-137 and miR-491 negatively regulate dopamine transporter expression and function in neural cells. *Neurosci. Bull.* 32 (6), 512–522. <https://doi.org/10.1007/s12264-016-0061-6>.
- Kalia, L.V., Lang, A.E., 2015. Parkinson's disease. *Lancet* 386 (9996), 896–912. [https://doi.org/10.1016/S0140-6736\(14\)61393-3](https://doi.org/10.1016/S0140-6736(14)61393-3).
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., Morishima, K., 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45 (D1), D353–D361. <https://doi.org/10.1093/nar/gkw1092>.
- Karpenko, M.N., Vasilishina, A.A., Gromova, E.A., Muruzheva, Z.M., Miliukhina, I.V., Bernadotte, A., 2018. Interleukin-1beta, interleukin-1 receptor antagonist, interleukin-6, interleukin-10, and tumor necrosis factor-alpha levels in CSF and serum in relation to the clinical diversity of Parkinson's disease. *Cell. Immunol.* 327, 77–82. <https://doi.org/10.1016/j.cellimm.2018.02.011>.
- Koper, O.M., Kaminska, J., Sawicki, K., Kemona, H., 2018. CXCL9, CXCL10, CXCL11, and their receptor (CXCR3) in neuroinflammation and neurodegeneration. *Adv. Clin. Exp. Med.* 27 (6), 849–856. <https://doi.org/10.17219/acem/68846>.
- Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J.P., Tamayo, P., 2015. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* 1 (6), 417–425. <https://doi.org/10.1016/j.cels.2015.12.004>.
- Li, J.H., Liu, S., Zhou, H., Qu, L.H., Yang, J.H., 2014. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data (Database issue), D92–97 *Nucleic Acids Res.* 42, 92–97. <https://doi.org/10.1093/nar/gkt1248>.
- Li, Y., Niu, M., Zhao, A., Kang, W., Chen, Z., Luo, N., Zhou, L., Zhu, X., Lu, L., Liu, J., 2019. CXCL12 is involved in  $\alpha$ -synuclein-triggered neuroinflammation of Parkinson's disease. *J. Neuroinflamm.* 16 (1), 263. <https://doi.org/10.1186/s12974-019-1646-6>.
- Mak, T.W., Grusdat, M., Duncan, G.S., Dostert, C., Nonnenmacher, Y., Cox, M., Binsfeld, C., Hao, Z., Brüstle, A., Itsumi, M., Jäger, C., Chen, Y., Pinkenburg, O., Camara, B., Ollert, M., Bindslev-Jensen, C., Vasilou, V., Gorrini, C., Lang, P.A., Lohoff, M., Harris, L.S., Hiller, K., Brenner, D., 2017. Glutathione primes T cell metabolism for inflammation. *Immunity* 46 (6), 1089–1090. <https://doi.org/10.1016/j.immuni.2017.06.009>.
- Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., Alizadeh, A.A., 2015. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. <https://doi.org/10.1038/nmeth.3337>.
- O'Sullivan, S.A., O'Sullivan, C., Healy, L.M., Dev, K.K., Sheridan, G.K., 2018. Sphingosine 1-phosphate receptors regulate TLR4-induced CXCL5 release from astrocytes and microglia. *J. Neurochem.* 144 (6), 736–747. <https://doi.org/10.1111/jnc.14313>.
- Oliveira Volpe, C.M., Vaz, T., Rocha-Silva, F., Villar-Delfino, P.H., Nogueira-Machado, J. A., 2020. Is galanin a promising therapeutic resource for neural and nonneural diseases? *Curr. Drug Targets* 21 (9), 922–929. <https://doi.org/10.2174/1389450121666200225112055>.
- Osterman, T.J., Terry, M., Miller, R.S., 2020. Improving cancer data interoperability: the promise of the minimal common oncology data elements (mCODE) initiative. *JCO Clin. Cancer Inform.* 4, 993–1001. <https://doi.org/10.1200/JCO.20.00059>.
- Prots, I., Winner, B., 2019. Th17 cells: a promising therapeutic target for Parkinson's disease? *Expert Opin. Ther. Targets* 23 (4), 309–314. <https://doi.org/10.1080/14728222.2019.1590336>.
- Ren, R., Sun, Y., Zhao, X., Pu, X., 2015. Recent advances in biomarkers for Parkinson's disease focusing on biochemicals, omics and neuroimaging. *Clin. Chem. Lab. Med.* 53 (10), 1495–1506. <https://doi.org/10.1515/cclm-2014-0783>.
- Ross, S.H., Cantrell, D.A., 2018. Signaling and function of interleukin-2 in T lymphocytes. *Annu. Rev. Immunol.* 36, 411–433. <https://doi.org/10.1146/annurev-immunol-042617-053352>.
- Sabatino Jr., J.J., Probstel, A.K., Zamvil, S.S., 2019. B cells in autoimmune and neurodegenerative central nervous system diseases. *Nat. Rev. Neurosci.* 20 (12), 728–745. <https://doi.org/10.1038/s41583-019-0233-2>.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. <https://doi.org/10.1101/gr.1239303>.
- Soreq, L., Guffanti, A., Salomonis, N., Simchovitz, A., Israel, Z., Bergman, H., Soreq, H., 2014. Long non-coding RNA and alternative splicing modulations in Parkinson's leukocytes identified by RNA sequencing. *PLoS Comput. Biol.* 10 (3), 1003517. <https://doi.org/10.1371/journal.pcbi.1003517>.
- Tan, J.S.Y., Chao, Y.X., Rotzschke, O., Tan, E.K., 2020. New insights into immune-mediated mechanisms in Parkinson's disease. *Int. J. Mol. Sci.* 21 (23) <https://doi.org/10.3390/ijms21239302>.
- Wang, R., Li, Q., He, Y., Yang, Y., Ma, Q., Li, C., 2020. miR-29c-3p inhibits microglial NLRP3 inflammasome activation by targeting NFAT5 in Parkinson's disease. *Genes Cells* 25 (6), 364–374. <https://doi.org/10.1111/gtc.12764>.
- Yeung, Y.T., Aziz, F., Guerrero-Castilla, A., Arguëlles, S., 2018. Signaling pathways in inflammation and anti-inflammatory therapies. *Curr. Pharm. Des.* 24 (14), 1449–1484. <https://doi.org/10.2174/1381612824666180327165604>.