

RESEARCH ARTICLE

Identification of biomarkers associated with inflammatory response in Parkinson's disease by bioinformatics and machine learning

Yatan Li¹✉, Wei Jia²✉, Chen Chen³, Cheng Chen⁴, Jinchao Chen³, Xinling Yang^{5,6*}, Pei Liu^{3*}

1 Department of Pharmacy, the Second Affiliated Hospital of Xinjiang Medical University, Urumqi, China, **2** The Affiliated Tumor Hospital of Xinjiang Medical University, Urumqi, China, **3** College of Information Science and Engineering, Xinjiang University, Urumqi, China, **4** College of Software, Xinjiang University, Urumqi, China, **5** Xinjiang Key Laboratory of Neurology, Urumqi, China, **6** Xinjiang Medical University, Urumqi, China

✉ These authors contributed equally to this work and should be considered co-first authors.

* yangxinling2014@163.com (XY); lp374187120@163.com (PL)



OPEN ACCESS

Citation: Li Y, Jia W, Chen C, Chen C, Chen J, Liu P, et al. (2025) Identification of biomarkers associated with inflammatory response in Parkinson's disease by bioinformatics and machine learning. PLoS One 20(5): e0320257. <https://doi.org/10.1371/journal.pone.0320257>

Editor: Naseem Ahamad, UT Health San Antonio: The University of Texas Health Science Center at San Antonio, UNITED STATES OF AMERICA

Received: June 5, 2024

Accepted: February 14, 2025

Published: May 28, 2025

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0320257>

Copyright: © 2025 Li et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/),

Abstract

Parkinson's disease (PD) is a common and debilitating neurodegenerative disorder. The inflammatory response is essential in the pathogenesis and progression of PD. The goal of this study is to combine bioinformatics and machine learning to screen for biomarker genes related to the inflammatory response in PD. First, differentially expressed genes associated with inflammatory response were screened, PPI networks were constructed and enriched for analysis. LASSO, SVM-RFE and Random Forest algorithms were used to screen biomarker genes. Then, ROC curves were drawn and PD risk predicting models were constructed on the basis of the biomarker genes. Finally, drug sensitivity analysis, mRNA-miRNA network construction and single-cell transcriptome data analysis were performed. The experimental results showed that we screened 31 differentially expressed genes related to inflammatory response. Signaling pathways such as cytokine activity were associated with these genes. Three biomarkers were identified using machine learning algorithms: IL18R1, NMUR1 and RELA. Seventeen co-associated miRNAs were identified by the mRNA-miRNA network as possible regulatory nodes in PD. Finally, these three biomarkers were found to be closely associated with T cells, Endothelial cells, excitatory neurons, inhibitory neurons, and other cells in single-cell transcriptomic analysis. In conclusion, IL18R1, NMUR1 and RELA could be potential therapeutic targets for PD in inflammatory response and new biomarkers for PD diagnosis.

1. Introduction

PD is an intractable and neurodegenerative disease associated with high morbidity and disability [1]. There is currently no effective treatment in the medical

which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The datasets generated and/or analyzed during the current study are available in the Gene Expression Omnibus (GEO) repository, under the accession number GSE20141, GSE20164 and GSE202210. The repository contains all relevant raw data needed to replicate the results of the study, including values for statistical analysis.

Funding: This study was funded by the Project about Parkinson's disease peripheral neuropathy drug intervention fusion research innovation team [2023D14006 to Yatan Li]. The funders participated in the study design, data collection and analysis, decision to publish, or manuscript writing.

Competing interests: The authors have declared that no competing interests exist.

community, and existing clinical treatments are ineffective [2]. The prevalence of PD is projected to continue to increase over the next 30 years, placing a significant financial and health care burden on society [3]. PD has a set of well-defined features, such as bradykinesia with resting tremor, as well as nonmotor symptoms including autonomic dysfunction, incontinence, and cognitive impairment [4]. When clinical signs of PD first appear, patients lose about 60% of striatal dopaminergic nerve endings and about 30% of nigrostriatal neurons [5]. Because the clinical features of Parkinson's disease are similar to those of other neurodegenerative diseases, and because there are no biomarkers for early and definitive diagnosis, misdiagnosis occurs even when the neurologist makes a diagnosis based on the patient's clinical manifestations and medication response. In addition, studies have shown that PD patients may experience these symptoms decades before diagnosis, resulting in patients not receiving timely treatment [6]. Therefore, the identification of more reliable and sensitive biomarkers is critical for early diagnosis and intervention in PD [7].

The initial manifestation of PD is an inflammatory response, and specific analysis of inflammation-associated genes may reveal useful biomarkers for monitoring disease progression [8]. Microglia-mediated neuroinflammation is thought to be an important aggravating factor driving dopaminergic neuron degeneration and central spread of pathogenic proteins. By exacerbating neuronal damage, neuroinflammation leads to neurodegeneration in PD [9]. Studying the underlying inflammatory mechanisms in PD may help to improve our understanding of disease progression. It may also aid in the development of immunomodulatory drugs and immunotherapeutic strategies to suppress inflammation. The neuroinflammation associated with the pathogenesis of PD has been widely studied. For example, activated microglia secrete a variety of pro-inflammatory cytokines, including tumor necrosis factor- α (TNF- α), interleukin (IL)-1 β , and interleukin IL-6 [10–12]. Heidari, A., et al. [13] concluded that the importance of neuroinflammation in PD patients is manifested by elevated serum and cerebrospinal fluid levels of pro-inflammatory cytokines and the appearance of active microglial cells in various regions of the CNS such as the substantia nigra. Araújo, B., et al. [14] showed that microglia remain hyperactivated over time, accompanied by a dysregulated secretion of pro-inflammatory cytokines, thus rendering microglia facilitators or contributing to the neuropathologic and toxic processes of PD. Lv, Q.-K., et al. [15] discussed the interplay of the microglia and α -syn in PD development and the possible mechanisms of microglial autoaggregation and phase phagocytosis in removing α -syn and suppressing neuroinflammation. Yi, M., et al. [16] found that several inflammatory genes are associated with PD risk, including genes related to neurotoxins, metabolism-related enzymes, and genes encoding cytokines. In summary, neuroinflammatory responses are highly implicated in the development and progression of PD. Analysis of inflammatory genes may reveal new PD markers and ideas for further understanding of PD pathogenesis.

In recent years, researchers have combined bioinformatics and machine learning to analyze data from samples such as blood and postmortem substantia nigra (SN)

from PD patients to search for potential biomarkers of PD and to assist clinicians in early diagnosis and detection of PD patients [17–20]. For example, Lei, C., et al. [17] searched for PD necrosis-associated genes by comprehensive cluster analysis, enrichment analysis, and WGCNA analysis. Yang, X. and Z. Wang [18] searched for new immune-related biomarkers of Parkinson's disease by developing a PD diagnostic model using Lasso and multivariate Cox regression. Chen, X., et al. [19] Search for potential biomarkers of Parkinson's disease using gene ontology (GO) enrichment and gene set variation analysis (GSVA). Bao, Y., et al. [20] screened genes closely related to PD using LASSO and SVM algorithms, which provided new perspectives for further research on PD immune mechanisms and treatment. In summary, the combination of bioinformatics technology and machine learning is expected to accurately identify potential biomarkers of PD, which is important for early detection of PD patients, patient stratification, and disease progression monitoring.

In this study, we downloaded the GSE20141 and GSE20164 datasets containing normal and PD samples from the gene expression omnibus(GEO) database platform [21]. Differential expression of inflammatory response genes was screened using the GSE20141 dataset. PPI networks, GO and KEGG pathway analysis were conducted for the differential expression genes using bioinformatics tools. Then Lasso, SVM-RFE and RSF machine learning algorithms were used to get candidate genes from these differential genes, and the intersection was taken to get the biomarkers for PD. Validate the diagnostic performance of biomarkers using ROC curves. Corresponding PD risk prediction models were also constructed to explore the predictive ability of these biomarkers. And the validity of these biomarkers was validated in the external dataset GSE20164. Subsequently, the biomarkers were subjected to drug-sensitive analysis and corresponding mRNA-miRNA regulatory network construction, which provided insight into the biomarkers' biological functions and potential mechanisms. Finally, single-cell transcriptomic analysis revealed that the screened biomarkers were abundantly expressed in different cell types. The entire analysis procedure is shown in Fig 1.

2. Materials and methods

2.1 Data collection

In this experiment, we obtained two PD-related datasets, GSE20141 and GSE20164, on GEO [22]. The biotypes for these datasets are all Homo sapiens and the sample tissues are from the substantia nigra of the brain. The GSE20141 database was used as the main study object, and the expression profiling data generated using the GPL570 data platform included 18 samples, of which 8 and 8 were in normal and PD group, respectfully. The GSE20164 database served as the external validation subject, and the GPL96 data platform was used to generate the expression profiling data, which contained 11 samples, of which 5 and 6 were in the control and PD groups, respectively. Two hundred genes related to the immune response were downloaded from the human gene set HALLMARK_INFLAMMATORY_RESPONSE by the Gene Set Enrichment Analysis (GSEA) platform [23].

2.2 Inflammation-related differential gene screening

Microarray data from the GSE20141 dataset were used to screen for differential gene expression between controls and PD groups with the R “limma” package [24]. Screening conditions for differential genes were $P < 0.05$, $|\log_2FC| > 1$. Heat-maps and volcano plots were used to illustrate these differential genes. The screened differential genes were taken to intersect with genes related to the inflammatory reaction.

2.3 Immune cell infiltration analysis

The proportion of 22 autoimmune tumor-infiltrating cell types in PD was evaluated by reverse convolution algorithm (CIBERSORT) [25]. Differential expression maps and correlation heat maps of immune cell components in the normal and PD groups were graphed with the appropriate R packages.

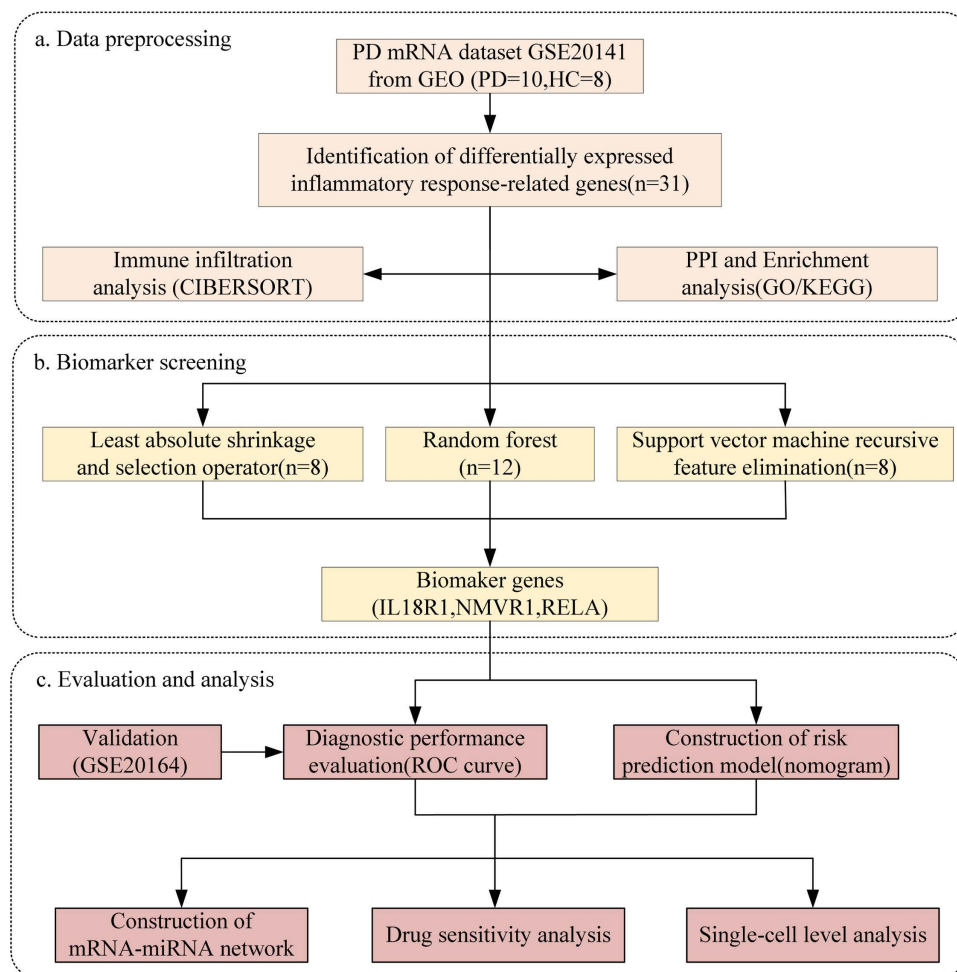


Fig 1. General workflow of this study.

<https://doi.org/10.1371/journal.pone.0320257.g001>

2.4 PPI network building and differential gene enrichment analysis

The differentially identified genes were input into the STRING online tool to obtain protein-protein interaction network (PPI) data, with the minimum interaction value set at 0.4. The PPI network of the screened differential genes was visualized in Cytoscape software [26]. Finally, differentially expressed genes associated with inflammatory response were analyzed for GO enrichment [27] and KEGG pathway enrichment [28], respectively. GO enrichment analysis can understand the common features of the screened differential genes in terms of biological processes, molecular functions, and cellular composition. KEGG pathway enrichment analysis explores the functions and regulatory mechanisms of the screened differential genes in organisms by detecting the enrichment of pathways in these genes.

2.5 Multiple machine learning algorithms for biomarker screening

The screened inflammation-related differential genes were further screened using LASSO [29], SVM-RFE [30], and RSF [31] algorithms. The LASSO regression algorithm is suitable for multicollinearity problems and is able to implement a regression method for variable selection along with parameter estimation. The complexity is adjusted by the parameter λ . For linear models with more variables, the penalty is greater for larger parameters. Finally, a model with fewer variables

can be obtained. Here, the optimal value of λ is continuously adjusted using the “glmnet” package of R. A backward selection method is the SVM-RFE algorithm. The RSF algorithm assigns an importance score to each differential gene, and feature screening is achieved by ranking the magnitude of the importance score of each gene. By further screening the previously screened differential genes using each of these machine learning algorithms, three gene sets were obtained. The intersection of these three gene sets is taken to obtain the potential biomarkers for PD.

2.6 Assessment of the diagnostic performance of biomarkers

On the GSE20141 and GSE20164 validation sets, we verified that these biomarkers have diagnostic value for PD using the ROC curve.

2.7 Biomarker-based PD risk prediction models

Using these screened biomarkers, we developed a biomarker gene-based risk prediction model for PD. The reliability of the risk predictive model for these biomarkers was verified by constructing a nomogram of PD patients, ROC curves for patient classification, calibration curves, and clinical curves.

2.8 Drug sensitivity analysis of biomarkers

We searched for biomarker-related drug information in the CellMiner database [32]. Finally, the results were visualized using the R package “ggplot2” [33].

2.9 Biomarker-based mRNA-miRNA regulatory networks

The biomarker-associated miRNAs were obtained using the online website Mirwalk [34]. This site is a comprehensive database of miRNA-target gene interactions containing information on miRNA target genes from a wide range of species. Interactions of biomarkers with relevant miRNAs are shown. The mRNA-miRNA regulatory network was mapped using the Cytoscape tool [26].

2.10 Processing and analyzing single-cell transcriptome data

The initial single-cell transcriptome data for Parkinson’s disease were taken from the GSE20228 dataset. During data processing, we utilized the R package Seurat (<http://satijalab.org/seurat/>) [35] to normalize, scale, and cluster the single-cell data. For quality control of the single cell data, $nFeature_RNA \leq 8000$, $nCount_RNA \leq 30000$, and $percent.mt \leq 3$ were used. Then, different cell types were tagged according to the specifically expressed genes in each cluster. Finally, the “runTSNE” function was used to visualize the distribution of different cell clusters.

3. Results

3.1 Identify differentially expressed genes related to PD inflammation

The PD dataset GSE20141 was screened for the differentially expressed genes using the method of limma of variance at $P < 0.05$ and $|\log_2FC| > 1$. As a result, 2855 differentially expression genes (DEGs) were identified. Among them, 394 were upregulated and 2461 were downregulated. These differentially expressed genes were visualized by heat map (Fig 2A). Then, intersections with 200 inflammation response related genes (IRRGs) were performed to yield 31 inflammation response related differentially expressed genes (Fig 2B). And from the volcano graphs of differentially expressed genes, it was found that 30 of these 31 differential expression genes related to inflammatory response were downregulated genes and 1 was upregulated gene (Fig 2C). Finally, the correlation between these 31 genes was analyzed. It was found that the PCDH7 gene showed negative correlation with most of the genes and positive correlation with some of the genes (Fig 2D).

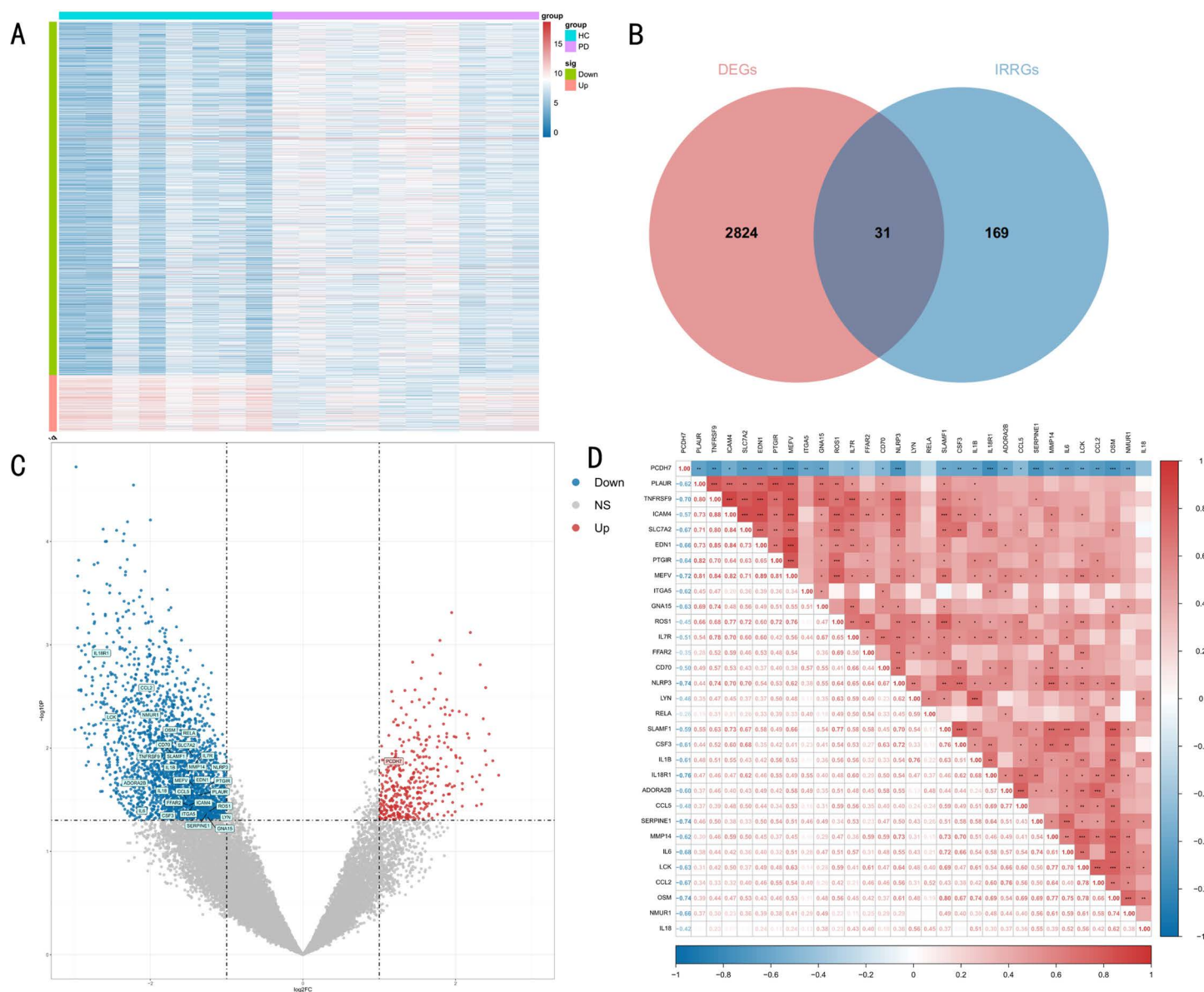


Fig 2. Results of the differentially regulated expression analysis of inflammatory response genes in PD. (A) Heat map of differential expression genes in PD. (B) Intersection Venn diagram of differential expression genes and inflammatory response related genes. (C) Volcano diagram of differential expression genes in PD. (D) Correlation diagram of 31 differentially expressed inflammatory response associated genes.

<https://doi.org/10.1371/journal.pone.0320257.g002>

3.2 Analysis of the correlation of immune cell infiltration in PD

Further analyses of 31 differentially expressing immune response genes examined the relationship between PD and immune-cell infiltrates. First, the proportions of the 22 different immune cells present in all samples of the GSE20141 dataset were calculated using the CIBERSORT algorithm (Fig 3A). The differential expression of immune cell infiltration in PD was then analyzed. The analysis showed differential expression of both activated mast cells and neutrophil in PD (Fig 3B). Then, the correlation between 22 immune cell infiltrates was analyzed again, and the results were shown in the correlation heat map (Fig 3C). Finally, association analysis was performed between differential expression of inflammation response

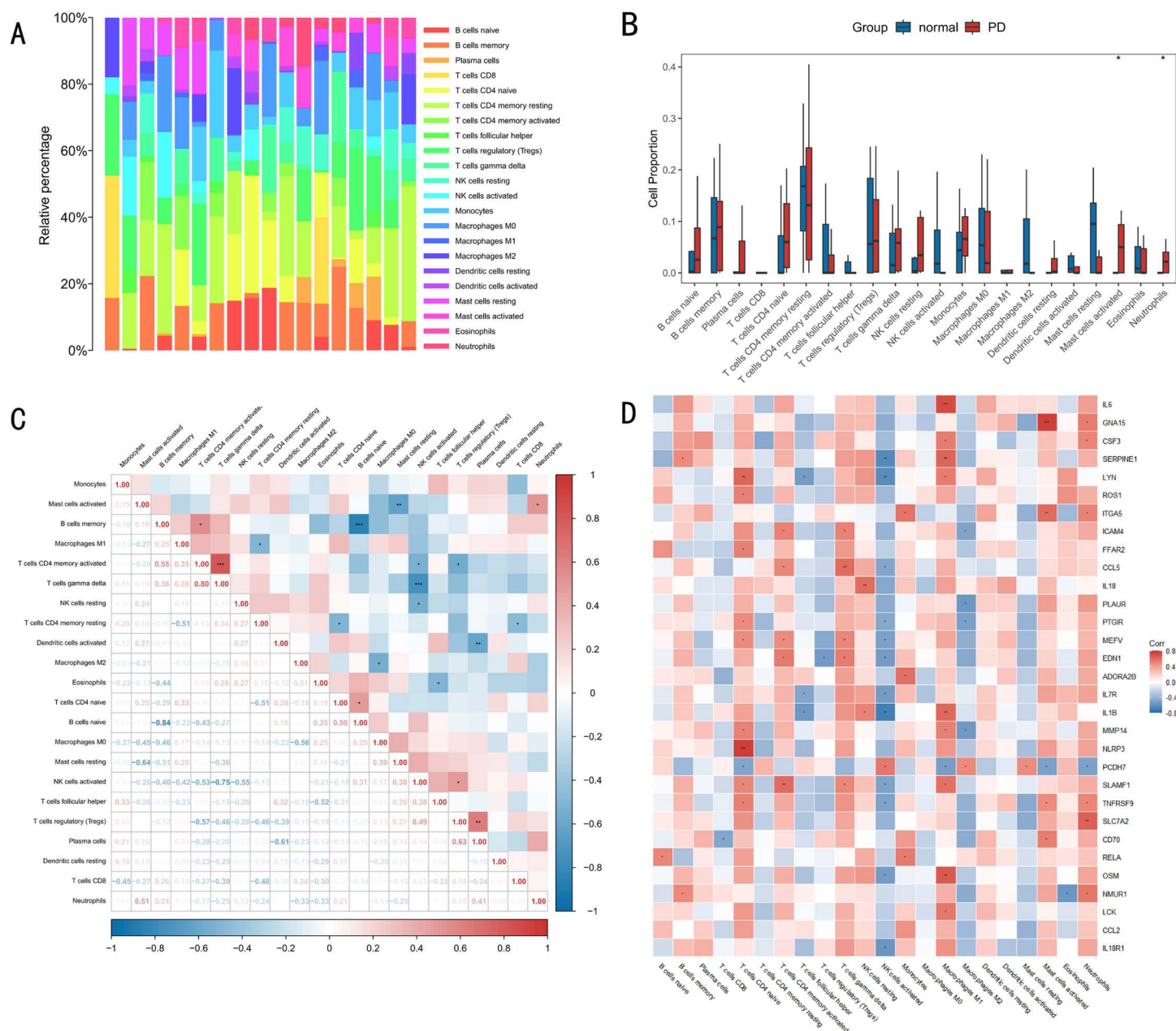


Fig 3. Results of correlation analysis of the immune cell infiltrates in PD. (A) Plots of 22 immune cell proportions in all samples of data set GSE20141. (B) Differential expression plot of the immune cell infiltrates. The blue and red colors indicate normal and PD samples, respectively. (C) Correlation plots of 22 immune-cell infiltrations. (D) Construction of association graph between differential expression of inflammatory response-related genes and immune-cell infiltrates. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

<https://doi.org/10.1371/journal.pone.0320257.g003>

related gene and immune cell infiltration (Fig 3D). The analysis showed a strong association between several differentially expressed inflammatory response related genes and immune cells, such as NLRP3 and CD4 naive T cells (Spearman $r = 0.77$, $p < 0.001$), GNA15 and activated mast cells (Spearman $r = 0.76$, $p < 0.001$), IL6 and M1 macrophages (Spearman $r = 0.72$, $p < 0.001$) showed a strong positive correlation.

3.3 Building PPI networks and analyzing GO and KEGG enrichment

In order to further investigate the possible interactions of the 31 differentially expressing inflammation response-related genes in PD, a PPI network was first constructed via the string website, and a sum of 31 nodes and 83 edges was obtained (Fig 4A). The PPI network was redrawn utilizing Cytoscape software. The shading of the gene color indicates the expression level (Fig 4B). To further investigate the function and localization of these 31 differentially expressed

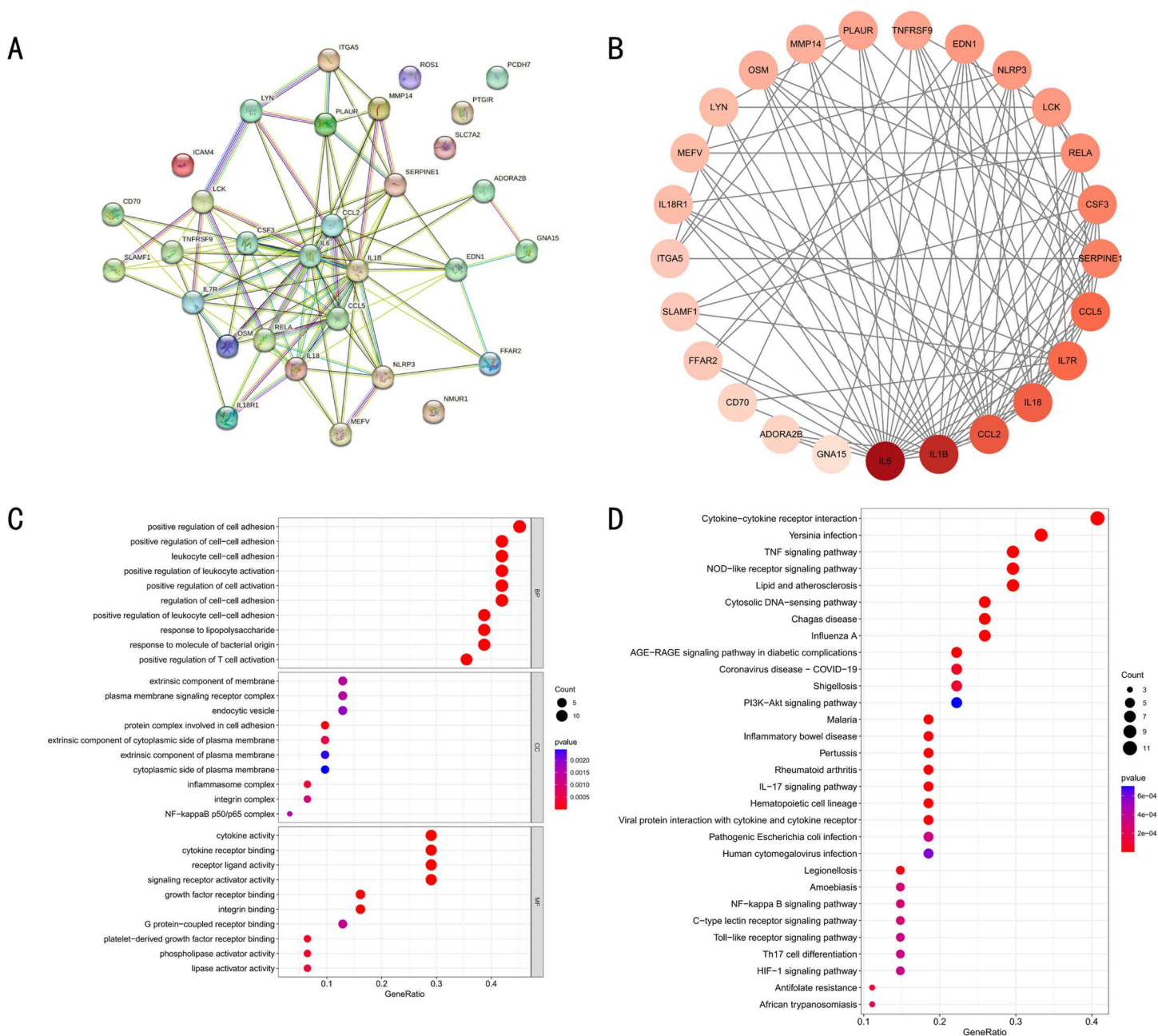


Fig 4. PPI network and results of GO and KEGG analysis. (A) PPI network based on 31 differentially expressed genes associated with inflammatory response. (B) PPI network plotted in Cytoscape software. (C) Graph of GO analysis results. (D) Graph of KEGG analysis results.

<https://doi.org/10.1371/journal.pone.0320257.g004>

inflammatory response-related genes, GO and KEGG analyses were conducted. GO enrichment analysis data indicated that most of these genes were associated with positive regulation of cell adhesion in the BP class and positive regulation of cell adhesion, etc., with extramembrane components and plasma membrane signaling receptor complexes, etc. in the CC class, and with cytokine activity and cytokine receptor binding, etc. in the MF class (Fig 4C). In addition, KEGG pathway analysis revealed that most of the genes were associated with cytokine-cytokine receptor interaction, Yersinia infection, and TNF signaling pathway (Fig 4D).

3.4 Machine learning-based biomarker screening

Three different machine learning algorithms were then applied to further screen the candidate genes. The 8-fold cross-validated LASSO algorithm was used first, and a total of eight candidate genes were screened (Fig 5A). They were IL18R1, NMUR1, RELA, CD70, SLC7A2, PTGIR, LYN, and GNA15. Next, 8-fold cross-validation was performed on the SVM-RFE algorithm. The false discovery rate was minimal when the candidate gene number was 8 (Fig 5B). They were RELA, CD70, NMUR1, SLAMF1, ADORA2B, IL18R1, CCL2, and IL6, respectively, and then the RandomForest algorithm was used to obtain the gene importance score ranking (Fig 5C). This RandomForest experiment was set to screen for candidate genes with MeanDecreaseGini>0.3. Thus, the number of candidate genes screened was 12, which were NMUR1, IL18R1, LCK, PTGIR, PCDH7, RELA, ITGA5, OSM, SLC7A2, EDN1, CCL2, and FFAR2. Finally, the candidate genes obtained from these algorithms are taken to intersect to obtain 3 biomarkers. The three biomarker genes were IL18R1, NMUR1, and RELA (Fig 5D).

3.5 Diagnostic performance assessment of biomarker genes

To further explore the potential diagnostic and prognostic properties of these three biomarkers in PD, this experiment first examined their expression in PD. The results showed that all three biomarker genes were expressed higher than normal in PD patients (Fig 6A). The diagnostic performance of three biomarkers, IL18R1, NMUR1 and RELA, was then evaluated. The ROC curves showed that IL18R1 (AUC:0.913), NMUR1 (AUC:0.838) and RELA (AUC:0.825) all had good diagnostic performance, suggesting that they may be potential therapeutic targets for PD (Fig 6B). The next step was to verify whether IL18R1, NMUR1, and RELA had generalization ability and robustness. This work was further validated on an external dataset, GSE20164, and found that NMUR1 and RELA were differentially expressed, while the IL18R1 differential expression profile was not significant (Fig 6C). Finally, the ROC curves showed that IL18R1 (AUC:0.633), NMUR1 (AUC:0.900), and RELA (AUC:0.867) also had good diagnostic performance on the validation set GSE20164 (Fig 6D).

3.6 Construction of a biomarker-based risk prediction model for PD

To further explore the risk predictive ability and clinical value of IL18R1, NMUR1 and RELA, this study integrated these three biomarkers and constructed a PD risk prediction model (Fig 7A). A score is obtained based on the high and low levels of expression of each biomarker gene in the nomogram, and the total score corresponds to its risk value. For example, patients with low expression of the RELA gene, high NMUR1 gene expression, and low IL18R1 gene expression would receive a total score of approximately 196 points, corresponding to a risk prediction probability of 0.5. Then, the ROC curve was plotted and the AUC of this model was 0.968, which indicates that the model has a strong prediction ability (Fig 7B). The calibration curve revealed a small difference between the real and the predicted PD risk, indicating the high prediction accuracy of the model (Fig 7C). Finally, the clinical impact curve showed that the model was more similar to the actual situation in clinical assessment (Fig 7D). Finally, it was verified that IL18R1, NMUR1 and RELA could be potential targets for PD prediction analysis with prognostic ability.

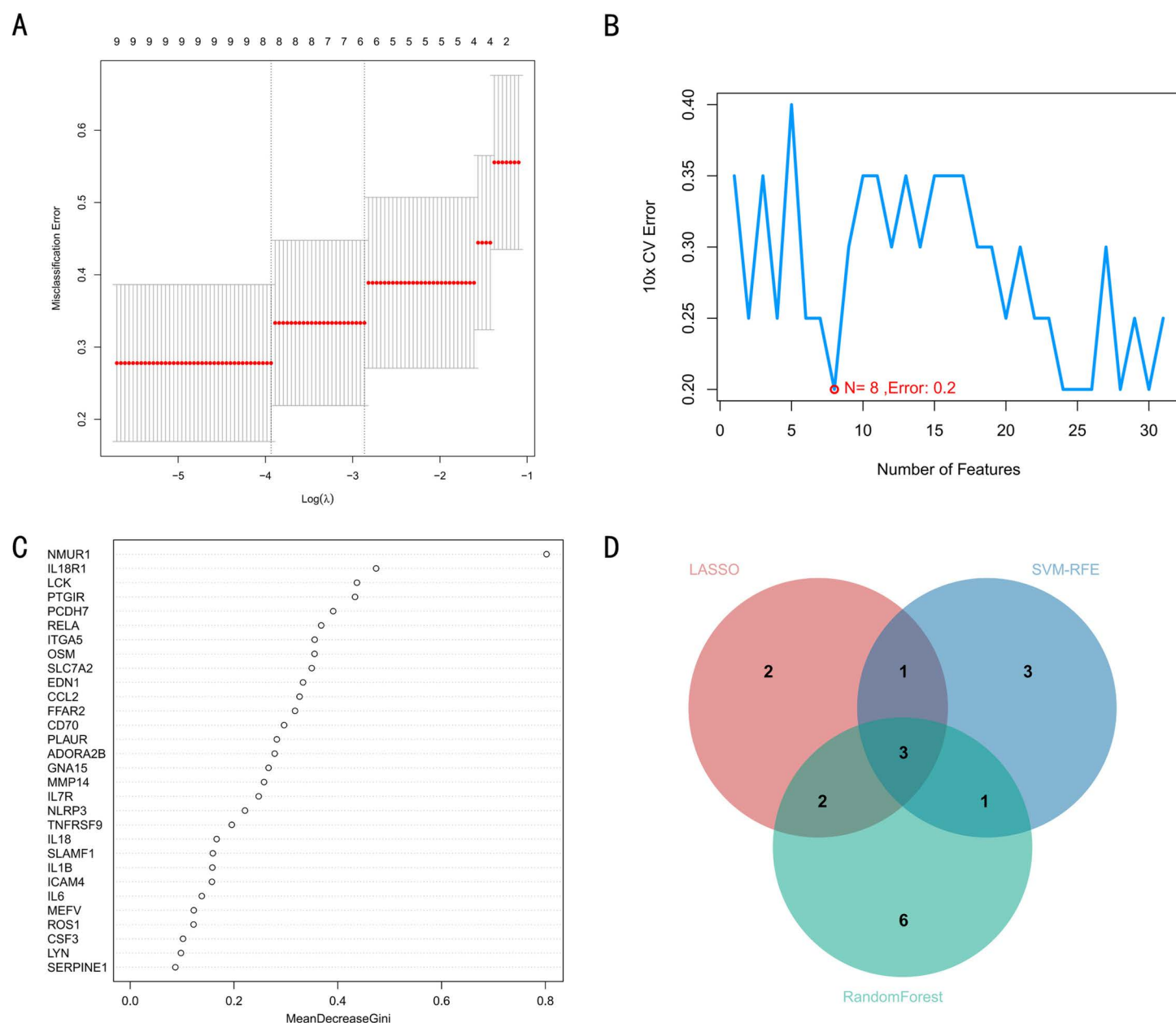


Fig 5. Results of machine learning based biomarker screening analysis. (A) Feature selection graph of LASSO algorithm. (B) Error rate graph of SVM-RFE algorithm. The red dots indicate the numbers of candidate candidate genes with the minimal false discovery rate. (C) Sorting graph of gene importance scores for the RandomForest algorithm. (D) Venn diagram of the intersection of the gene candidates examined by the three machine learning algorithms.

<https://doi.org/10.1371/journal.pone.0320257.g005>

3.7 Drug sensitivity analysis of biomarkers

Drug sensitivity analysis helps identify potential drug targets for Parkinson's patients and generates new ideas for treating Parkinson's patients. The CellMiner database was used to correlate the drug sensitivity of IL18R1, NMUR1 and RELA. In this paper, $P < 0.05$ and $|\text{Cor}| > 0.3$ were used as screening criteria. IL18R1 was positively correlated with the

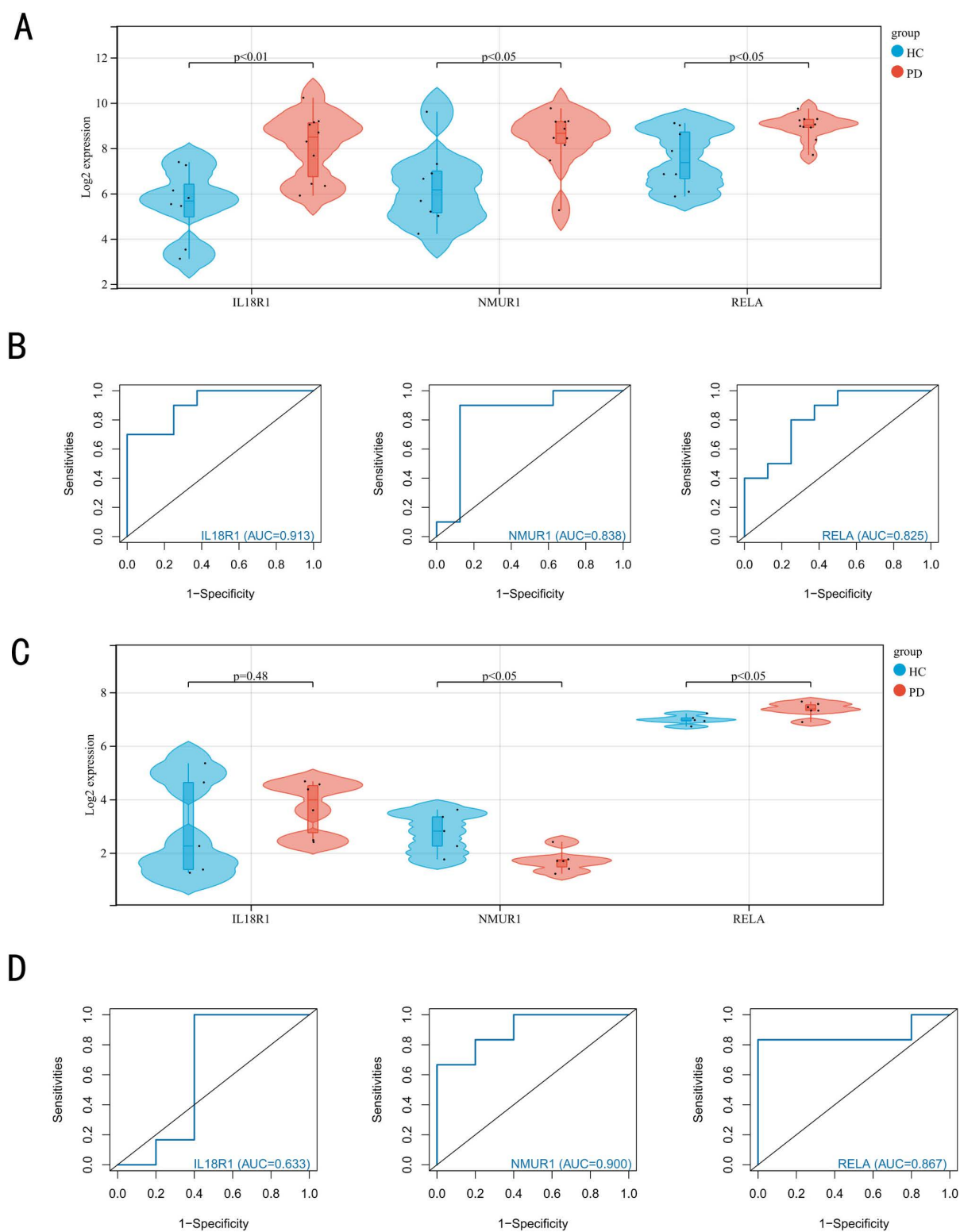


Fig 6. Expression and diagnostic performance analysis results of IL18R1, NMUR1 and RELA genes. (A) Expression of biomarker genes in GSE20141 dataset. (B) Evaluation of biomarkers with ROC curves. (C) Expression of biomarker genes on validation set GSE20164. (D) ROC curve evaluating the diagnostic performance of biomarker genes on validation set GSE20164.

<https://doi.org/10.1371/journal.pone.0320257.g006>

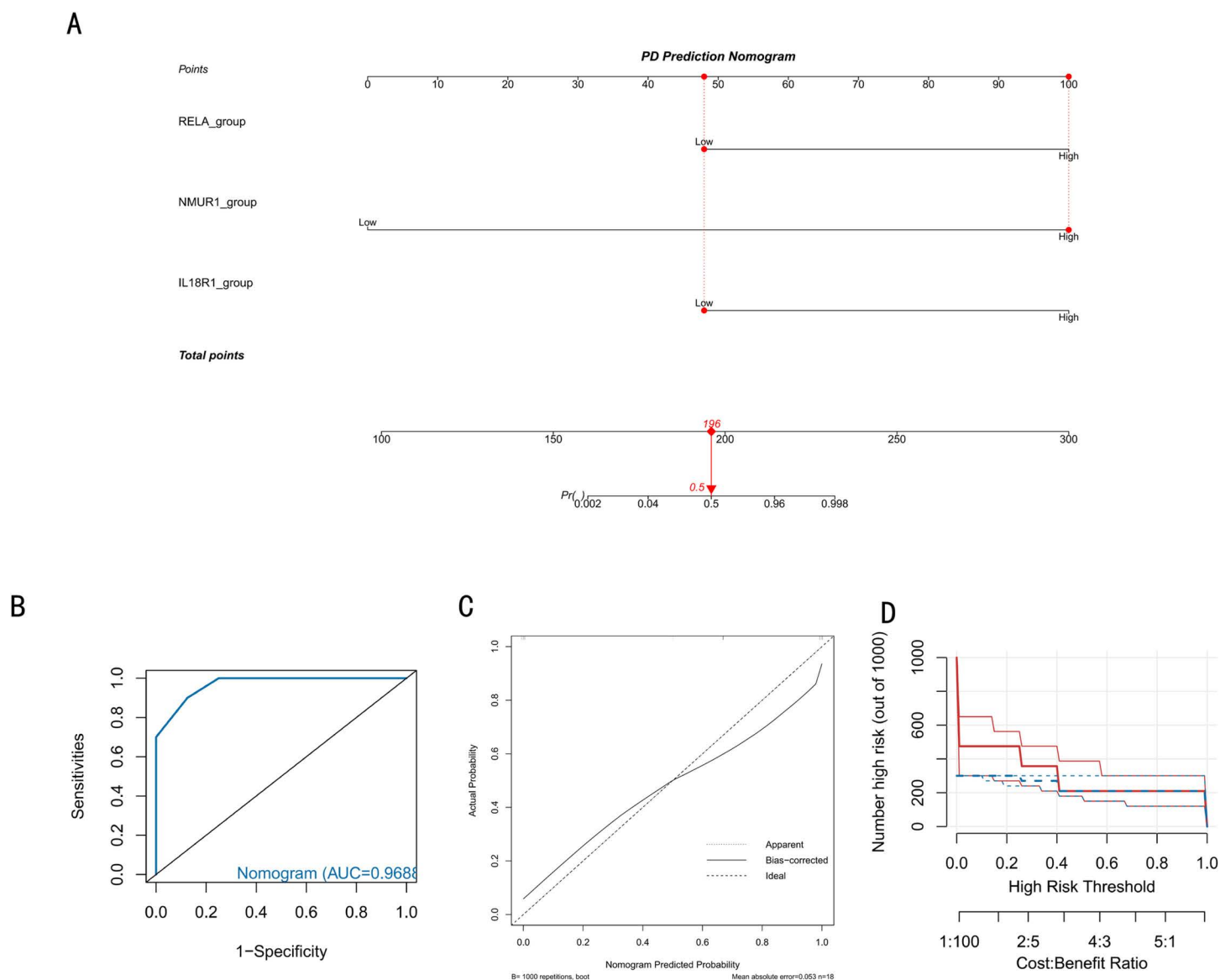


Fig 7. Construction of PD risk predicting model and evaluation indices. (A) Construction of the nomogram of the PD disease risk prognosis model on the basis of the three biomarkers. (B) ROC curve of the model. (C) The calibration curve evaluating the predictive accuracy of the model. (D) Clinical impact curve evaluating the model's clinical performance.

<https://doi.org/10.1371/journal.pone.0320257.g007>

drugs Alisertib (Cor=0.31, P=0.016), Azacitidine (Cor=0.33, P=0.01) and Motesanib (Cor=0.35, P=0.007), IL18R1 was negatively correlated with the drug BP-1–82 (Cor=-0.36, P=0.004), NMUR1 was negatively correlated with Lexibulin (Cor=0.31, P=0.017) and Norvir (Cor=0.48, P<0.001), RELA was positively correlated with CAMPTOTHECIN (Cor=0.32, P=0.012) and Clofarbine (Cor=0.3, P=0.019) and 7 other drugs, and RELA was positively correlated with ARRY-162 (Cor=-0.37, P=0.004) and Tamoxifen (Cor=-0.3, P=0.018) and 7 other drugs (Fig 8).

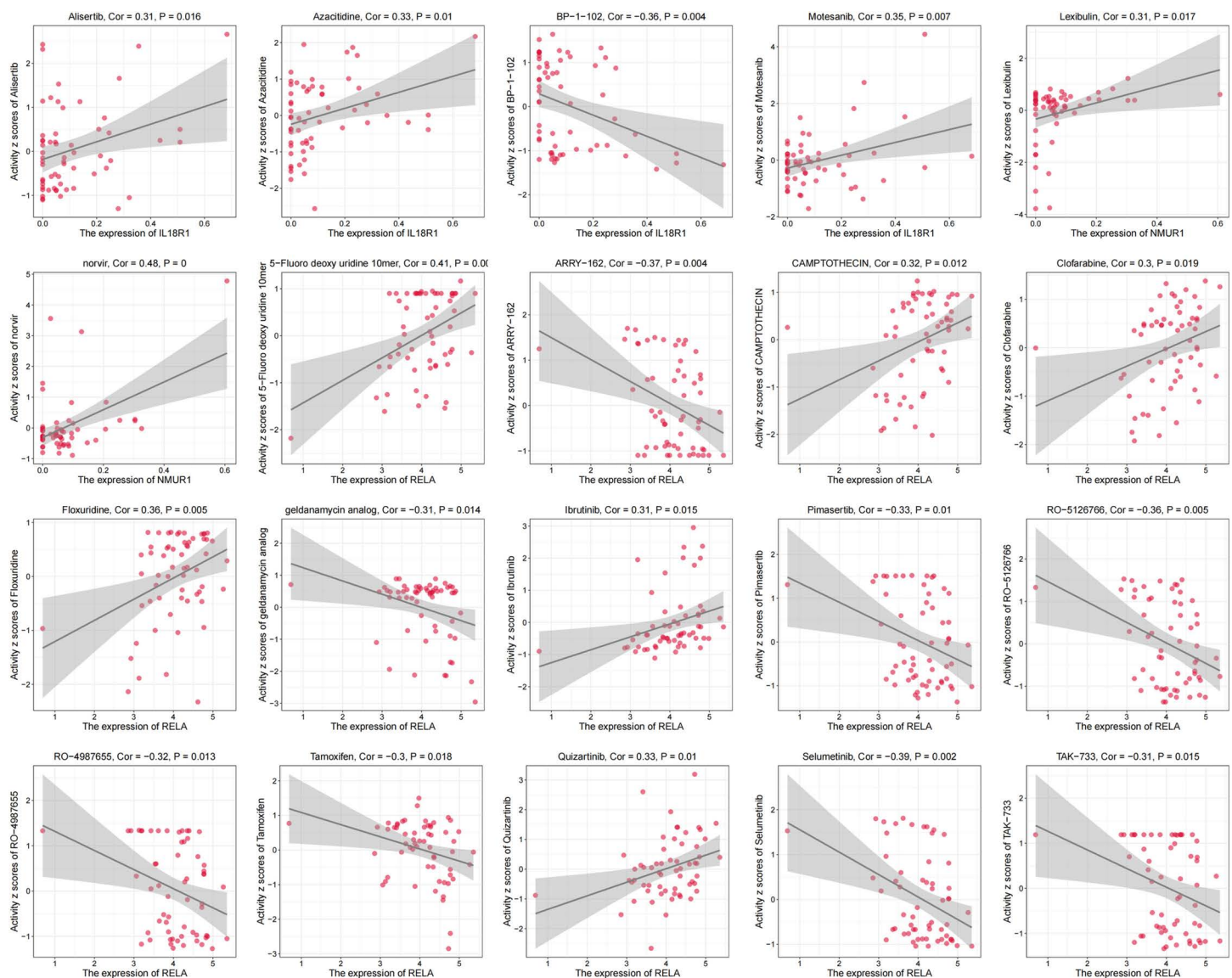


Fig 8. Results of drug sensitivity correlation analysis based on IL18R1, NMUR1 and RELA genes.

<https://doi.org/10.1371/journal.pone.0320257.g008>

3.8 Construction of a biomarker-based mRNA-miRNA network

To further investigate the upstream regulatory networks and potential molecular mechanisms of the three biomarker genes IL18R1, NMUR1 and RELA. In this paper, we queried the miRNAs upstream of the three biomarker gene via the Mirwalk online database, and the results showed that there were 427 miRNAs associated with IL18R1 gene, 325 miRNAs associated with NMUR1 gene, and 374 miRNAs associated with RELA gene. And Venn diagram was drawn to screen 17 miRNAs with common predictive relationships of the three biomarker genes (Fig 9A). Then, the mRNA-miRNA regulatory networks of the three biomarker genes IL18R1, NMUR1 and RELA with the 17 miRNAs were mapped using Cytoscape software (Fig 9B). We further explored the roles of these 17 miRNA target genes in the relevant pathways (S1 Fig). Through KEGG and GO analysis of miRNA target genes, we found that the pathways of miRNA target genes overlapped

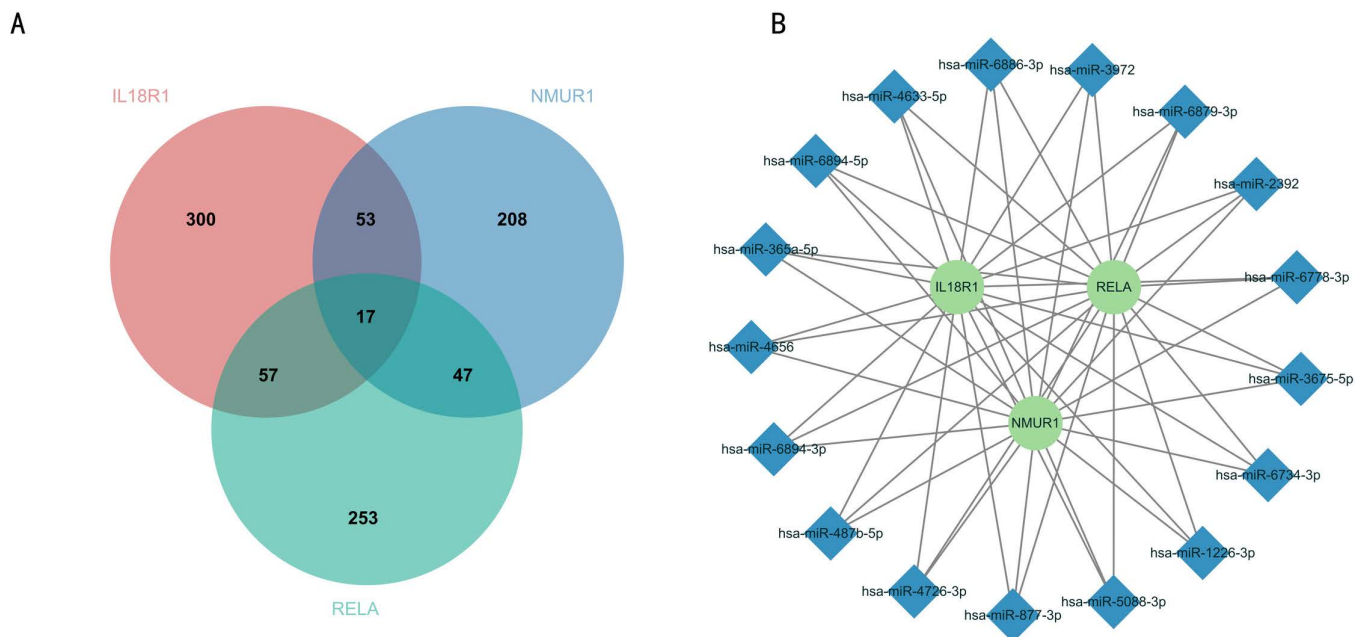


Fig 9. Results of mRNA-miRNA network construction based on IL18R1, NMUR1 and RELA biomarkers. (A) The Venn diagram shows the miRNAs associated with IL18R1, NMUR1, and RELA. (B) Mapping of the mRNA-miRNA regulatory network of IL18R1, NMUR1, and RELA.

<https://doi.org/10.1371/journal.pone.0320257.g009>

with those at the mRNA level. These pathways, although significant ($P < 0.05$), were ranked low and therefore did not appear in the pathway heatmap for mRNAs. Specifically, these include the virus process and ECM-receptor interaction pathways. The final results indicated that these 17 miRNAs could be potential regulatory targets in PD, and the network constructed with IL18R1, NMUR1 and RELA biomarkers was one of the important regulatory networks in PD.

3.9 Single-cell transcriptomic data biomarker expression

Clustering and manually annotating the raw single-cell transcriptome data yielded eight major cell types (Fig 10A). All of these cells are intimately involved in the pathogenesis of PD. One example is microglia, the most abundant immune cell in the central nervous system, which have been linked to the progression of PD. And to further understand the relative proportions of specific cells in the total cells, we plotted stacked bars for each sample (Fig 10B). In further detail, we displayed the expression of IL18R1, NMUR1, and RELA in these cell populations in bubble plots (Fig 10C) and feature plots (Fig 10D) to further reveal the role of these biomarkers in the progression of PD. The results indicated that RELA was abundantly expressed in these cell populations. IL18R1 and NMUR1 were abundantly expressed in endothelial cells and inhibitory neurons, respectively, and were closely associated with other cell populations.

4. Discussion

PD is a progressive neurological disease characterized by the death of midbrain dopaminergic neurons, α -synuclein aggregates, and movement disorders [36]. A major factor in the decline of dopaminergic neurons is the neuroinflammatory response [37]. Because PD is directly linked to genetics, single nucleotide polymorphisms (SNPs) in immune and inflammatory genes may increase the risks of developing PD in people with PD by affecting the body's immune system and inflammation response [16].

In this work, we initially identify 2855 differently represented genes in the GSE20141 dataset. We used bioinformatics tools to do this. Then, 31 inflammatory response-related differently expressing genes were obtained by intersecting with 200

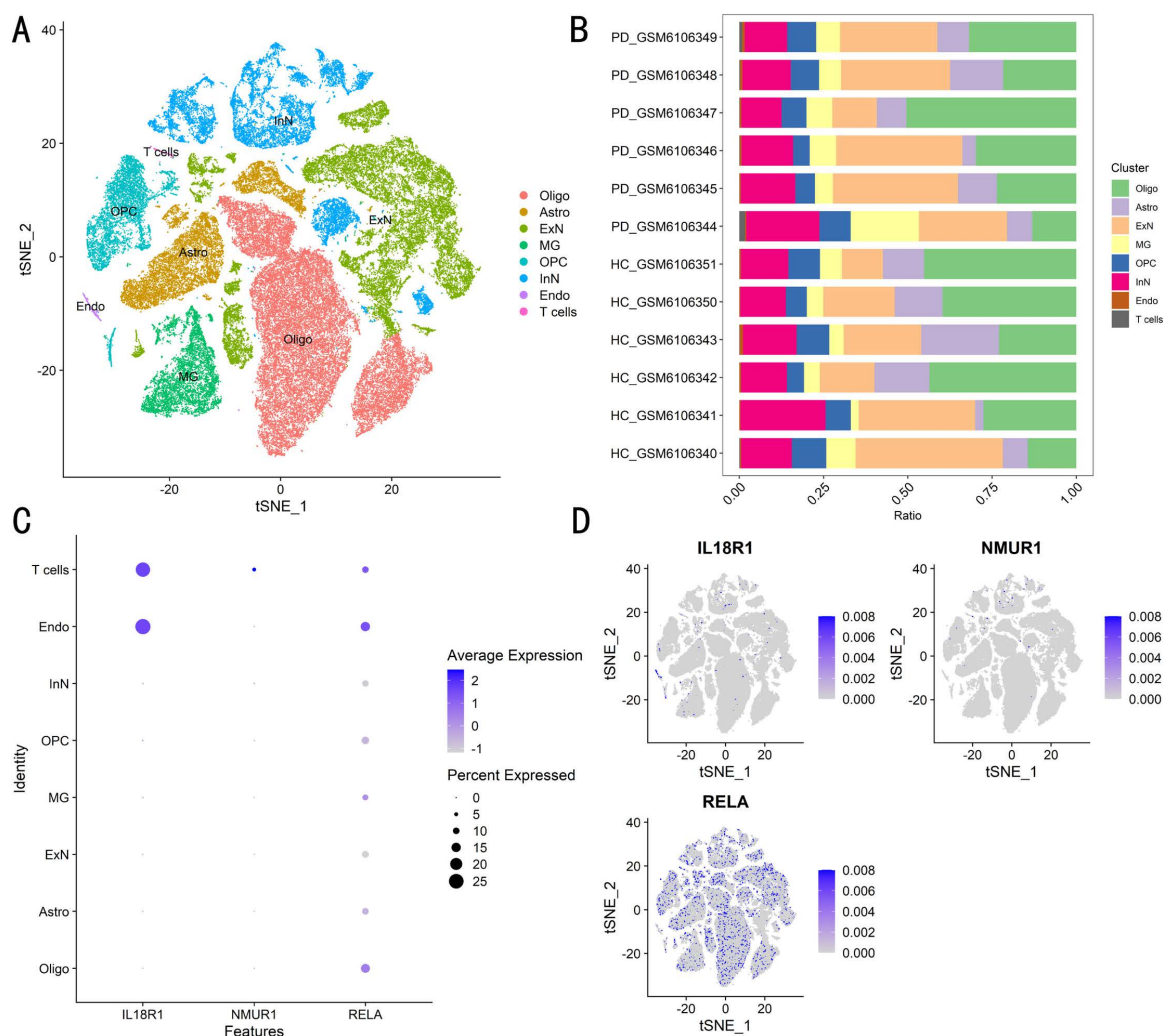


Fig 10. Analysis of IL18R1, NMUR1 and RELA based single cell transcriptome data. (A) The 8 major cell types detected in the single cell transcriptome dataset GSE20228. (B) The proportions of different cell types in the total cells. (C) Bubble plots of IL18R1, NMUR1 and RELA expression in these cell populations. (D) Characterization plots of IL18R1, NMUR1 and RELA expression in these cell populations.

<https://doi.org/10.1371/journal.pone.0320257.g010>

human inflammatory response corresponding genes. And these 31 inflammatory response-related differentially expressed genes were evaluated by PPI network, GO analysis and KEGG pathway analysis. Second, LASSO, SVM-REF and RSF machine learning algorithms were employed to further screen the features, and three biomarkers, IL18R1, NMUR1 and RELA, were obtained, and these three biomarkers were validated in the external dataset GSE20164. IL18R1, NMUR1 and RELA were subjected to drug sensitivity analysis and these three biomarkers were found to be closely associated with several drug components. Specifically, IL18R1 was negatively correlated with the drug BP-1–102 (Cor=-0.36, P=0.004), and RELA was negatively correlated with the drugs ARRY-162 (Cor=-0.37, P=0.004) and tamoxifen (Cor=-0.3, P=0.018). The negative correlation implies that the effect of the drug tends to diminish when the expression level of the gene is elevated. For example, Xu, F., et al. [38] found a negative correlation between the expression of SLC1A5 and the sensitivity to camptothecin-based drugs in their analysis of drug sensitivity to classical chemotherapeutic agents. Additionally, because this study is in the preliminary exploratory stage, drugs with weak correlations may still uncover potential complex biological

mechanisms. The relationship between gene expression and drug response is often nonlinear, so a single weak correlation does not necessarily indicate a lack of biological significance. For example, Song, D., et al. [39] found that the drug Sapitinib, which has a weak correlation with NRAV expression, can induce apoptosis and inhibit the phosphorylation of the epidermal growth factor receptor (EGFR) and its downstream pathways when investigating the sensitivity of small-molecule drugs related to NRAV expression. 17 co-associated miRNAs were identified by the mRNA-miRNA network, which may be one of the regulatory nodes in PD. In single-cell transcriptomic analysis, the three biomarkers screened were abundantly expressed in different cell populations. IL18R1 was specifically expressed in T cells and Endothelial cells. It suggests that IL18R1 may play an important role in the immune response, especially in T cells. T cells are key cells of the immune system and are involved in recognizing and attacking infected or tumor cells [40]. Whereas IL18R1 is a subunit of the IL-18 receptor, IL-18 is a pro-inflammatory cytokine that is commonly associated with immune response and inflammation [41]. RELA is expressed in almost all cell populations. The expression of RELA in virtually all cell populations suggests that it is a fundamental and broadly regulated factor that may maintain essential functions and responsiveness in the physiological state of the cell, especially in response to inflammation, infection, or cellular damage [42]. NMUR1 is specifically expressed in T cells. NMUR1 is usually associated with neuropeptide signaling, which may influence the immune response by regulating T cell activation, proliferation, or function [43]. It is worth noting that although the three biomarkers we screened showed extremely strong correlations with PD in ROC curves, nomograms, and various types of bioconfidence analyses on external datasets. However, all current related studies only indirectly indicate that these three biomarkers and PD have an association. Therefore, as newly discovered biomarkers for PD, their reliability warrants further investigation.

The protein encoded by IL18R1 is a member of the interleukin-1 (IL-1) receptor family of cytokine receptors [44]. IL18R1 binds specific to interleukin 18 (IL18) and has an important function in IL18-mediated signaling. Prior studies have demonstrated that IL-18 is a key mediator of neuroinflammation and neurodegeneration in the central nervous system under pathological conditions [45]. The most prominent PPI hotspot identified by Henderson, A.R., et al. [46] in their analysis was the IL18/IL18R1 pathway, which clearly demonstrated the up-regulatory effect of hypomethylation on IL18/IL18R1 expression and highlighted the contribution of IL18 to the neuroinflammatory process in PD. In summary, the IL18R1 biomarker screened acts as an inflammatory gene that may influence the progression of Parkinson's disease.

Neuromedin U contains 2 receptors (NMUR1 and NMUR2) that are distributed throughout the nervous system and many tissues and organs, and has a variety of biological roles [47]. NMUR1 and NMUR2 have multiple physiological functions in the peripheral tissue and the central nervous system (CNS) [48]. The transcription factor nuclear factor- κ B (NF- κ B), a key regulator of inflammation and apoptosis, has been implicated in the pathology of various neural degenerative diseases [49]. The NF- κ B family is made up of five distinct subunits that bind to each other to form transcriptionally active dimers [50]. Nigrostriatal dopamine (DA) neurons in patients with PD have abnormally increased nuclear levels of RELA (NF- κ B subunit) [51]. NF- κ B is implicated in processes such as inflammation and immunity and acts as a pleiotropic transcription factor. Choi, D.C., et al. [52] found that NF- κ B inhibition, but not its activity, is the underlying cell death mechanisms in MPP+ toxicity. Lanzillotta, A., et al. [53] found that the acetylated state of RELA subunits promotes both neuroprotection and neurotoxicity. Overall, disease progression in PD patients is accompanied by the development of an inflammatory response, and there is a potential relationship between the inflammatory gene RELA and the pathogenesis of PD patients.

In conclusion, three biomarker genes, IL18R1, NMUR1 and RELA, identified in this paper are strongly associated with PD. These three genes may be potential therapeutic targets for PD with prognostic predictive value and clinical value as potential biomarker genes for PD.

5. Conclusion

Although the three biomarker genes we investigated showed high correlation with PD in ROC curve analysis on outside datasets, drug susceptibility, mRNA-miRNA regulatory networks, single-cell transcriptome data analysis, and related

research reports. However, the specific roles of the three identified biomarkers in PD need to be elucidated by further clinical and animal studies. However, drug sensitivity analyses were not performed on single-cell transcriptomics data as the primary objective of this study was to validate the specific expression of marker genes in different cell types rather than to explore the direct effects of drugs on these cell types. However, in the future, we will focus on drug sensitivity analysis of single-cell transcriptomics data.

In conclusion, in this study we identified three PD-related inflammatory response genes IL18R1, NMUR1 and RELA using a combination of several bioinformatics methods and several machine learning algorithms. These biomarkers have an enormous impact on the monitoring of PD progression. They can help clinicians detect the progression of PD earlier, enable more accurate diagnosis, and improve the prognosis of PD patients.

Supporting information

S1 Fig. S1 Fig KEGG and GO analysis of miRNA target genes.

(TIF)

S1 File. Data Availability Statement.

(DOCX)

Author contributions

Conceptualization: Yatan Li, Wei Jia, Chen Chen, Cheng Chen, Jinchao Chen, Pei Liu, Xinling Yang.

Data curation: Yatan Li, Wei Jia, Chen Chen, Cheng Chen, Jinchao Chen, Pei Liu, Xinling Yang.

Formal analysis: Yatan Li, Wei Jia, Chen Chen, Cheng Chen, Jinchao Chen, Pei Liu, Xinling Yang.

Investigation: Yatan Li, Chen Chen, Cheng Chen, Jinchao Chen, Pei Liu, Xinling Yang.

Methodology: Yatan Li, Wei Jia.

Project administration: Yatan Li, Wei Jia.

Resources: Yatan Li, Wei Jia.

Software: Yatan Li, Wei Jia.

Supervision: Yatan Li, Wei Jia.

Validation: Yatan Li, Wei Jia.

Visualization: Yatan Li, Wei Jia.

Writing – original draft: Yatan Li.

Writing – review & editing: Yatan Li, Wei Jia.

References

1. Tansey MG, Wallings RL, Houser MC, Herrick MK, Keating CE, Joers V. Inflammation and immune dysfunction in Parkinson disease. *Nat Rev Immunol.* 2022;22(11):657–73. <https://doi.org/10.1038/s41577-022-00684-6> PMID: [35246670](#)
2. Chopade P, Chopade N, Zhao Z, Mitragotri S, Liao R, Chandran Suja V. Alzheimer's and Parkinson's disease therapies in the clinic. *Bioeng Transl Med.* 2023;8(1):e10367.
3. Tolosa E, Garrido A, Scholz SW, Poewe W. Challenges in the diagnosis of Parkinson's disease. *Lancet Neurol.* 2021;20(5):385–97. [https://doi.org/10.1016/S1474-4422\(21\)00030-2](https://doi.org/10.1016/S1474-4422(21)00030-2) PMID: [33894193](#)
4. Mehanna R, Smilowska K, Fleisher J, Post B, Hatano T, Pimentel Piemonte ME, et al. Age cutoff for early-onset parkinson's disease: recommendations from the International Parkinson and Movement Disorder Society task force on early onset Parkinson's Disease. *Mov Disord Clin Pract.* 2022;9(7):869–78. <https://doi.org/10.1002/mdc3.13523> PMID: [36247919](#)
5. Yazdani U, Zaman S, Hynan LS, Brown LS, Dewey RB, Karp D, et al. Blood biomarker for Parkinson disease: peptoids. *npj Parkinson's Dis.* 2016;2(1):1–7.

6. Lindqvist D, Kaufman E, Brundin L, Hall S, Surova Y, Hansson O. Non-motor symptoms in patients with Parkinson's disease—correlations with inflammatory cytokines in serum. *J Neuroinflammation*. 2012.
7. Magalhães P, Lashuel HA. Opportunities and challenges of alpha-synuclein as a potential biomarker for Parkinson's disease and other synucleinopathies. *npj Parkinson's Dis*. 2022;8(1):93.
8. Pajares M, Rojo AI, Manda G, Boscal L, Cuadrado A. Inflammation in Parkinson's disease: mechanisms and therapeutic implications. *Cells*. 2020;9(7):1687.
9. Liu S-H, Wang Y-L, Jiang S-M, Wan X-J, Yan J-H, Liu C-F. Identifying the hub gene and immune infiltration of Parkinson's disease using bioinformatical methods. *Brain Res*. 2022;1785:147879. <https://doi.org/10.1016/j.brainres.2022.147879> PMID: 35278479
10. Williams-Gray CH, Wijeyekoon R, Yarnall AJ, Lawson RA, Breen DP, Evans JR, et al. Serum immune markers and disease progression in an incident Parkinson's disease cohort (ICICLE-PD). *Mov Disord*. 2016;31(7):995–1003. <https://doi.org/10.1002/mds.26563> PMID: 26999434
11. Hofmann KW, Schuh AFS, Saute J, Townsend R, Fricke D, Leke R, et al. Interleukin-6 serum levels in patients with Parkinson's disease. *Neurochem Res*. 2009;34(8):1401–4. <https://doi.org/10.1007/s11064-009-9921-z> PMID: 19214748
12. Scalzo P, Kümmer A, Cardoso F, Teixeira AL. Increased serum levels of soluble tumor necrosis factor- α receptor-1 in patients with Parkinson's disease. *J Neuroimmunol*. 2009;216(1–2):122–5.
13. Heidari A, Yazdanpanah N, Rezaei N. The role of Toll-like receptors and neuroinflammation in Parkinson's disease. *J Neuroinflammation*. 2022;19(1):135. <https://doi.org/10.1186/s12974-022-02496-w> PMID: 35668422
14. Araújo B, Caridade-Silva R, Soares-Guedes C, Martins-Macedo J, Gomes ED, Monteiro S, et al. Neuroinflammation and Parkinson's disease—from neurodegeneration to therapeutic opportunities. *Cells*. 2022;11(18):2908. <https://doi.org/10.3390/cells11182908> PMID: 36139483
15. Lv Q, Tao K, Wang X, Yao X, Pang M, Liu J, et al. Role of α -synuclein in microglia: autophagy and phagocytosis balance neuroinflammation in Parkinson's disease. *Inflamm Res*. 2023;72(3):443–62.
16. Yi M, Li J, Jian S, Li B, Huang Z, Shu L, et al. Quantitative and causal analysis for inflammatory genes and the risk of Parkinson's disease. *Front Immunol*. 2023;14:1119315. <https://doi.org/10.3389/fimmu.2023.1119315> PMID: 36926335
17. Lei C, Zhongyan Z, Wenting S, Jing Z, Liyun Q, Hongyi H, et al. Identification of necroptosis-related genes in Parkinson's disease by integrated bioinformatics analysis and experimental validation. *Front Neurosci*. 2023;17:1097293. <https://doi.org/10.3389/fnins.2023.1097293> PMID: 37284660
18. Yang X, Wang Z. Identification of novel immune-related biomarker and therapeutic drugs in Parkinson disease via integrated bioinformatics analysis. *Medicine (Baltimore)*. 2023;102(31):e34456. <https://doi.org/10.1097/MD.00000000000034456> PMID: 37543820
19. Chen X, Cao W, Zhuang Y, Chen S, Li X. Integrative analysis of potential biomarkers and immune cell infiltration in Parkinson's disease. *Brain Res Bull*. 2021;177:53–63. <https://doi.org/10.1016/j.brainresbull.2021.09.010> PMID: 34536521
20. Bao Y, Wang L, Yu F, Yang J, Huang D. Parkinson's disease gene biomarkers screened by the LASSO and SVM algorithms. *Brain Sci*. 2023;13(2):175. <https://doi.org/10.3390/brainsci13020175> PMID: 36831718
21. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002;30(1):207–10. <https://doi.org/10.1093/nar/30.1.207> PMID: 11752295
22. Zheng B, Liao Z, Locascio JJ, Lesniak KA, Roderick SS, Watt ML, et al. PGC-1 α , a potential therapeutic target for early intervention in Parkinson's disease. *Sci Transl Med*. 2010;2(52):52ra73. <https://doi.org/10.1126/scitranslmed.3001059> PMID: 20926834
23. Mootha VK, Lindgren CM, Eriksson K-F, Subramanian A, Sihag S, Lehar J, et al. PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*. 2003;34(3):267–73. <https://doi.org/10.1038/ng1180> PMID: 12808457
24. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. <https://doi.org/10.1093/nar/gkv007> PMID: 25605792
25. Li J, Zhang Y, Lu T, Liang R, Wu Z, Liu M, et al. Identification of diagnostic genes for both Alzheimer's disease and Metabolic syndrome by the machine learning algorithm. *Front Immunol*. 2022;13:1037318. <https://doi.org/10.3389/fimmu.2022.1037318> PMID: 36405716
26. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang P-L, Lotia S, et al. A travel guide to Cytoscape plugins. *Nat Methods*. 2012;9(11):1069–76. <https://doi.org/10.1038/nmeth.2212> PMID: 23132118
27. Mubeen S, Tom Kodamullil A, Hofmann-Apitius M, Domingo-Fernández D. On the influence of several factors on pathway enrichment analysis. *Brief Bioinform*. 2022;23(3):bbac143. <https://doi.org/10.1093/bib/bbac143> PMID: 35453140
28. Huang X, Hu H, Liu J, Zhang X, Jiang Y, Lv L, et al. Immune analysis and small molecule drug prediction of hepatocellular carcinoma based on single sample gene set enrichment analysis. *Cell Biochem Biophys*. 2022;80(2):427–34. <https://doi.org/10.1007/s12013-022-01070-8> PMID: 35195822
29. Qu X, Zhao X, Lin K, Wang N, Li X, Li S, et al. M2-like tumor-associated macrophage-related biomarkers to construct a novel prognostic signature, reveal the immune landscape, and screen drugs in hepatocellular carcinoma. *Front Immunol*. 2022;13:994019. <https://doi.org/10.3389/fimmu.2022.994019> PMID: 36177006
30. Li L, Ching W-K, Liu Z-P. Robust biomarker screening from gene expression data by stable machine learning-recursive feature elimination methods. *Comput Biol Chem*. 2022;100:107747. <https://doi.org/10.1016/j.compbiolchem.2022.107747> PMID: 35932551

31. Toprak B, Brandt S, Brederecke J, Gianfagna F, Vishram-Nielsen JKK, Ojeda FM, et al. Exploring the incremental utility of circulating biomarkers for robust risk prediction of incident atrial fibrillation in European cohorts using regressions and modern machine learning methods. *Europace*. 2023;25(3):812–9. <https://doi.org/10.1093/europace/euac260> PMID: 36610061
32. Reinhold WC, Sunshine M, Liu H, Varma S, Kohn KW, Morris J, et al. CellMiner: a web-based suite of genomic and pharmacologic tools to explore transcript and drug patterns in the NCI-60 cell line set. *Cancer Res*. 2012;72(14):3499–511. <https://doi.org/10.1158/0008-5472.CAN-12-1370> PMID: 22802077
33. Wickham H. Ggplot2. Wiley Interdiscip Rev Comput Stat. 2011;3(2):180–5.
34. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: an online resource for prediction of microRNA binding sites. *PLoS One*. 2018;13(10):e0206239. <https://doi.org/10.1371/journal.pone.0206239> PMID: 30335862
35. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36(5):411–20. <https://doi.org/10.1038/nbt.4096> PMID: 29608179
36. Cabrera Randaldi EDLRM, Nuytemans K, Martinez A, Luca CC, Keane RW, de Rivero Vaccari JP. Proof-of-principle study of inflammasome signaling proteins as diagnostic biomarkers of the inflammatory response in Parkinson's disease. *Pharmaceuticals (Basel)*. 2023;16(6):883. <https://doi.org/10.3390/ph16060883> PMID: 37375830
37. Luk KC, Kehm V, Carroll J, Zhang B, O'Brien P, Trojanowski JQ, et al. Pathological α -synuclein transmission initiates Parkinson-like neurodegeneration in nontransgenic mice. *Science*. 2012;338(6109):949–53. <https://doi.org/10.1126/science.1227157> PMID: 23161999
38. Xu F, Wang H, Pei H, Zhang Z, Liu L, Tang L, et al. SLC1A5 prefers to play as an accomplice rather than an opponent in pancreatic adenocarcinoma. *Front Cell Dev Biol*. 2022;10:800925. <https://doi.org/10.3389/fcell.2022.800925> PMID: 35419359
39. Song D, Wang X, Wang Y, Liang W, Luo J, Zheng J, et al. Integrated analysis of N1-methyladenosine methylation regulators-related lncRNAs in hepatocellular carcinoma. *Cancers (Basel)*. 2023;15(6):1800. <https://doi.org/10.3390/cancers15061800> PMID: 36980686
40. Ruf B, Greten TF, Korangy F. Innate lymphoid cells and innate-like T cells in cancer—at the crossroads of innate and adaptive immunity. *Nat Rev Cancer*. 2023;23(6):351–71.
41. Di Y, Wang Z, Xiao J, Zhang X, Ye L, Wen X, et al. ACSL6-activated IL-18R1-NF- κ B promotes IL-18-mediated tumor immune evasion and tumor progression. *Sci Adv*. 2024;10(38):eadp0719. <https://doi.org/10.1126/sciadv.adp0719> PMID: 39292786
42. Khan A, Zhang Y, Ma N, Shi J, Hou Y. NF- κ B role on tumor proliferation, migration, invasion and immune escape. *Cancer Gene Ther*. 2024;1–12.
43. Ye Y, Liang Z, Xue L. Neuromedin U: potential roles in immunity and inflammation. *Immunology*. 2021;162(1):17–29. <https://doi.org/10.1111/imm.13257> PMID: 32888314
44. Dale M, Nicklin MJ. Interleukin-1 receptor cluster: gene organization of IL1R2, IL1R1, IL1RL2 (IL-1Rrp2), IL1RL1 (T1/ST2), and IL18R1 (IL-1Rrp) on human chromosome 2q. *Genomics*. 1999;57(1):177–9. <https://doi.org/10.1006/geno.1999.5767> PMID: 10191101
45. Felderhoff-Mueser U, Schmidt OI, Oberholzer A, Bührer C, Stahel PF. IL-18: a key player in neuroinflammation and neurodegeneration? *Trends Neurosci*. 2005;28(9):487–93.
46. Henderson AR, Wang Q, Meechoovet B, Siniard AL, Naymik M, De Both M, et al. DNA methylation and expression profiles of whole blood in Parkinson's disease. *Front Genet*. 2021;12:640266. <https://doi.org/10.3389/fgene.2021.640266> PMID: 33981329
47. You C, Zhang Y, Xu P, Huang S, Yin W, Eric Xu H, et al. Structural insights into the peptide selectivity and activation of human neuromedin U receptors. *Nat Commun*. 2022;13(1):2045. <https://doi.org/10.1038/s41467-022-29683-w> PMID: 35440625
48. Aiyar N, Disa J, Foley JJ, Buckley PT, Wixted WE, Pullen M, et al. Radioligand binding and functional characterization of recombinant human NmU1 and NmU2 receptors stably expressed in clonal human embryonic kidney-293 cells. *Pharmacology*. 2004;72(1):33–41. <https://doi.org/10.1159/000078630> PMID: 15292653
49. Bhakar AL, Tannis L-L, Zeindler C, Russo MP, Jobin C, Park DS, et al. Constitutive nuclear factor-kappa B activity is required for central neuron survival. *J Neurosci*. 2002;22(19):8466–75. <https://doi.org/10.1523/JNEUROSCI.22-19-08466.2002> PMID: 12351721
50. Oeckinghaus A, Ghosh S. The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harb Perspect Biol*. 2009;1(4):a000034. <https://doi.org/10.1101/cshperspect.a000034> PMID: 20066092
51. Bellucci A, Bubacco L, Longhena F, Parrella E, Faustini G, Porri V, et al. Nuclear factor- κ B dysregulation and α -synuclein pathology: critical interplay in the pathogenesis of Parkinson's disease. *Front Aging Neurosci*. 2020;12:68. <https://doi.org/10.3389/fnagi.2020.00068> PMID: 32265684
52. Choi DC, Chae Y-J, Kabaria S, Chaudhuri AD, Jain MR, Li H, et al. MicroRNA-7 protects against 1-methyl-4-phenylpyridinium-induced cell death by targeting RelA. *J Neurosci*. 2014;34(38):12725–37. <https://doi.org/10.1523/JNEUROSCI.0985-14.2014> PMID: 25232110
53. Lanzillotta A, Sarnico I, Ingrassia R, Boroni F, Branca C, Benarese M, et al. The acetylation of RelA in Lys310 dictates the NF- κ B-dependent response in post-ischemic injury. *Cell Death Dis*. 2010;1(11):e96. <https://doi.org/10.1038/cddis.2010.76> PMID: 21368872