Neuropsychiatric Disease and Treatment

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Uncovering Potential Biomarkers and Constructing a Prediction Model Associated with Iron Metabolism in Parkinson's Disease

Yan Cheng[1,2], Hongjiang Zhai[2], Yong Liu[2], Yunzhou Yang[2], Bo Fang[2], Mingxiang Song[2], Ping Zhong[1]

[1]Department of Neurology, Suzhou Hospital of Anhui Medical University, Suzhou, Anhui, People's Republic of China; [2]Department of Neurology, Lu'an Hospital of Anhui Medical University, Lu'an, Anhui, People's Republic of China

Correspondence: Ping Zhong, Department of Neurology, Suzhou Hospital of Anhui Medical University, 616 Bianyang Third Road, Suzhou City, Anhui Province, 234000, People's Republic of China, Tel +86-18269816229, Email smile692427@tzc.edu.cn

Background: Parkinson's disease (PD) is a common neurodegenerative disorder. Iron metabolism abnormalities have been reported in PD patients and may contribute to disease pathogenesis. Our study aimed to explore key genes associated with iron metabolism in PD patients.

Methods: Three datasets and iron metabolism-related genes (IMRGs) were collected from the public database, and the datasets were merged into a combined dataset. PD-related differentially expressed genes (DEGs) were obtained and intersected with IMRGs to acquire iron metabolism-related DEGs (IMRDEGs). Subsequently, the IMRDEGs were subjected to functional enrichment and ROC analyses. Finally, key genes were identified, followed by the construction and evaluation of a risk score model, drug prediction, and RT-qPCR analysis.

Results: A total of 24 IMRDEGs were obtained. The AUC values of the 24 IMRDEGs ranged from 0.599 to 0.781. After logistic regression and the SVM analyses, a total of 10 key genes were identified, followed by the construction of the risk score model. The AUC value of the risk score model was 0.953, demonstrating good diagnostic value. The calibration curve and decision curve analysis showed that the risk score model has good predictive performance and clinical benefit for PD patients. Additionally, a total of 49 drugs were predicted.

Conclusion: A total of 10 key genes were identified as potential biomarkers, and the risk score model was constructed for PD patients, exhibiting good diagnostic. This study may provide potential biomarkers for PD patients, promoting an understanding of the pathogenesis of PD.

Keywords: Parkinson's disease, iron metabolism, risk score model, diagnosis

Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disease worldwide.[1] The primary pathological features of PD include reduced dopamine production in the substantia nigra (SN), degeneration and loss of dopaminergic neurons, abnormal α-synuclein (α-syn) aggregation, and mitochondrial dysfunction.[2] Clinical features of PD mainly involve resting tremors, muscle rigidity, bradykinesia, and loss of postural reflexes.[3] Moreover, PD is also associated with depression.[4,5] In 2021, approximately 508,400 new PD cases and 5.08 million patients were reported in China, with incidence and prevalence rates of 35.7 and 356.8 per 100,000, respectively.[6] Globally, the number of PD patients is projected to reach 13 million cases by 2040, according to the Global Burden of Disease Study.[7]

The diagnosis of PD primarily relies on clinical symptom identification.[8] However, diagnostic accuracy remains suboptimal due to overlapping symptoms with other neurodegenerative disorders and the lack of definitive biomarkers for early detection.[9] Misdiagnosis rates in clinical practice range from 15% to 24%.[10,11] While neuroimaging and some biochemical markers show promise, they are often costly, invasive, or lack specificity.[9] Therefore, there is an urgent need for reliable, non-invasive biomarkers to facilitate early diagnosis, monitor disease progression, and potentially predict PD risk.

Received: 30 December 2024
Accepted: 24 February 2025
Published: 28 February 2025

Accumulating evidence suggests a significant link between iron metabolism dysregulation and PD pathogenesis.[12] Iron plays a crucial role in various physiological processes in the central nervous system, including oxidative stress, production of ATP, processes of myelination and remyelination, and neurotransmitter synthesis and metabolism.[13] In PD patients, abnormal iron accumulation, particularly in the SN, contributes to oxidative stress, α-syn aggregation, and neuronal death, all of which are key features of PD pathology.[14] Therefore, the association between iron metabolism and PD should not be overlooked.

Most of the previous studies only stayed at the observation level, and the specific regulatory network and molecular mechanism of iron metabolism-related genes in the pathogenesis of PD were not deeply explored. Although it is known that abnormal iron accumulation can lead to pathological features of PD, there is still no clear conclusion on which genes play a core regulatory role in this process and how they interact with each other. Moreover, there is also a lack of effective risk prediction models to accurately assess an individual's risk of PD based on iron metabolism related indicators. This study is aimed at these research gaps. Through bioinformatics analysis, key genes related to iron metabolism were identified, and their potential mechanisms in PD were explored. Constructing a risk scoring model based on key genes may provide a new tool for individual PD risk assessment and fill the gap in the field of risk prediction. Drug prediction of key genes is expected to discover new therapeutic targets and potential drugs, provide new ideas and methods for the clinical treatment of PD, and promote the transformation from basic research to clinical application.

Materials and Methods

Data Collection and Preprocessing

The workflow of this study is presented in Figure 1. A total of three PD-related datasets, GSE6613, GSE54536, and GSE22491, were obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo) using the "GEOquery" package in R.[15] All samples were obtained from human blood. Specifically, the GSE6613 (platform: GPL96) dataset includes 50 PD samples and 22 healthy controls,[16] the GSE54536 (platform: GPL10558) dataset contains 4 PD samples and 4 healthy controls,[17] and the GSE22491 (platform: GPL6480) dataset contains 10 PD samples and 8 healthy controls.[18] Iron metabolism-related genes (IMRGs) were obtained from the GeneCards database (https://www.genecards.org) and the Yao et al study.[19]

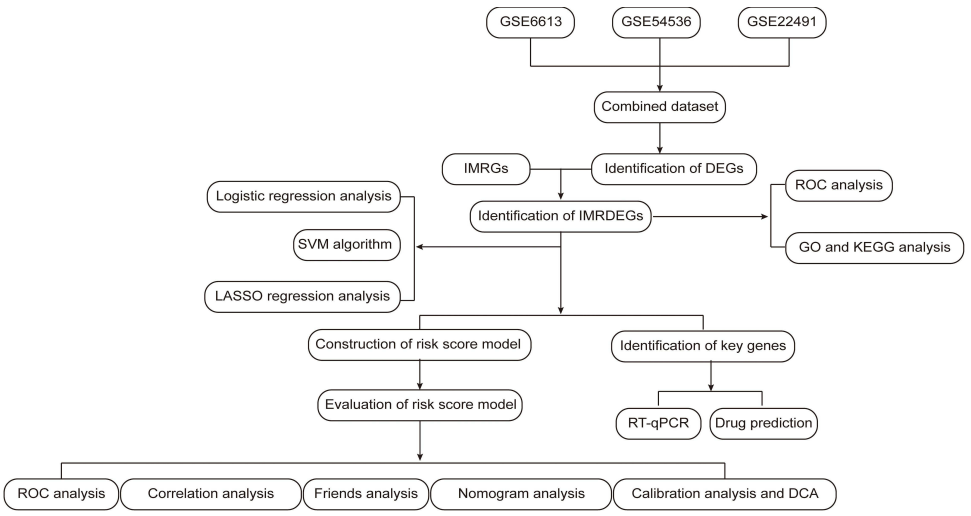

Figure 1 Flowchart of the study.

Abbreviations: IMRGs, iron metabolism-related genes; DEGs, differentially expressed genes; IMRDEGs, iron metabolism-related DEGs; ROC, receiver operating characteristic; GO, gene ontology; KEGG, Kyoto encyclopedia of genes and genomes; LASSO, least absolute shrinkage and selection operator; SVM, support vector machine; DCA, decision curve analysis.

After removing the batch effect across the three datasets using the "sva" package in R, the combined dataset was obtained.[20] The "limma" package in R was utilized for normalization.[21] Principal component analysis (PCA) was conducted to validate the elimination of batch effects.[22]

Identification of Iron Metabolism-Related Differentially Expressed Genes (DEGs) in PD

PD-related DEGs with $|\log_2$ fold change (FC)$| > 0.25$ and p-value < 0.05 in the PD patients were acquired in the combined dataset using the "limma" package in R.[21] The volcano plot was used to visualize the PD-related DEGs using the "ggplot2" package in R. Then, the PD-related DEGs were intersected with IMRGs to obtain iron metabolism-related differential expressed genes (IMRDEGs). Venn diagram was used to show the intersection of gene set, and heatmap was used to display the expression profiles of IMRDEGs. The location of IMRDEGs in chromosomes was employed through the "RCircos" package in R.[23]

Functional Enrichment Analysis

The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were performed to investigate the functional roles of IMRDEGs using the "clusterProfiler" package in R.[24] The screening criteria were set as adjusted p-value < 0.05 and false discovery rate (FDR, also known as q-value) < 0.25. The Benjamini-Hochberg method was used for multiple testing corrections.

Receiver Operating Characteristic (ROC) Analysis

To evaluate the diagnostic performance of IMRDEGs, the ROC analysis was performed using the "pROC" package in R.[25] The area under the curve (AUC) was used to measure the performance of the models.

Construction of the Risk Score Model

Logistic regression analysis was first performed on IMRDEG, and genes with p-value < 0.05 were selected, which were visualized using a forest plot. Subsequently, the selected genes were further refined using the support vector machine (SVM) algorithm to determine the optimal number of genes with minimal error rate and the highest accuracy.[26] Lastly, the least absolute shrinkage and selection operator (LASSO) regression analysis was applied to the selected genes in the SVM model using the "glmnet" package in R, and the results were visualized through the LASSO regression model and variable trajectory plots.[27] The IMRDEGs identified in the risk score model were defined as key genes. The risk score calculation formula is as follows:

$$\text{Risk } Score = \sum_i Coefficient\ (gene_i) \times mRNA\ Expression\ (gene_i)$$

Evaluation of the Risk Score Model

The diagnostic value of the risk score was assessed through ROC analysis via the "pROC" package in R. Then, the correlation analysis among key genes was carried out in the combined dataset using the Spearman correlation analysis, with results visualized by heatmap and scatter plot. Additionally, the Friends analysis was performed to explore the relationships among key genes using the "GOSemSim" package in R.[28] A nomogram was generated to visualize the risk score model through the "rms" package in R.[29] Moreover, a calibration curve was plotted to assess the accuracy of the model. The utility of the risk score model for decision-making was evaluated using decision curve analysis (DCA) via the "ggDCA" package in R.[30]

Drug Prediction

The Comparative Toxicogenomics Database (CTD, https://ctdbase.org) was utilized to predict drugs for key genes.[31] Drugs with a reference count >1 were selected. Finally, an mRNA-drug regulatory network was constructed using Cytoscape software.

Sample Collection and RT-qPCR

This study was approved by the Ethics Committee of Suzhou Hospital of Anhui Medical University (2024LL028) and written informed consent was obtained from all individuals. A total of 19 blood samples from 9 PD patients and 10 healthy controls were collected. PD patients were diagnosed by at least three experienced movement disorder neurologists trained in research, following the UK Brain Bank criteria for PD. Exclusion criteria included: a history of deep brain stimulation or cancer treatment, severe depression, dementia, liver or kidney disease, stroke or other cerebrovascular diseases, and other neurological disorders (such as Alzheimer's disease, Huntington's disease, etc). Healthy subjects had no family or personal history of neurodegenerative diseases. Subsequently, RNA extraction, reverse transcription, and RT-qPCR were performed using the blood RNA extraction kit (Magen, R4163-02), FastQuant cDNA first strand synthesis kit (TIANGEN, KR106), and SuperReal PreMix Plus (SYBR Green) (TIANGEN, FP205). The relative quantitative analysis of data was carried out by $2^{-\Delta\Delta CT}$ method. GAPDH and ACTB were the reference genes in RT-qPCR.

Statistical Analysis

All data processing and analyses were performed using the R software (version 4.3.0). Comparisons between the two groups were performed using the Wilcoxon rank-sum test. Unless otherwise specified, Spearman correlation analysis was used to evaluate the correlation. P-value <0.05 was considered statistically significant.

Results

Identification of IMRDEGs

A total of 93 IMRGs were obtained from the GeneCards database and 295 from the Yao et al study. After removing duplicates, 363 IMRGs were obtained (Table S1). After removing the batch effects, a total of 98 samples were acquired, including 64 PD samples and 34 healthy controls (Supplementary Figure 1A-D). Then, a total of 1145 PD-related DEGs, 596 up- and 549 down-regulated, were obtained in the combined dataset (Figure 2A). After intersecting with 363 IMRGs, 24 IMRDEGs were identified (Figure 2B). The heatmap of 24 IMRDEGs is shown in Figure 2C. The chromosomal locations of the 24 IMRDEGs are presented in Figure 2D.

Functional Enrichment Analysis

To elucidate the biological functions of the 24 IMRDEGs, GO and KEGG enrichment analyses were carried out. GO analysis of biological processes showed that these genes were involved in erythrocyte homeostasis, erythrocyte differentiation, myeloid cell homeostasis, response to oxidative stress, and regulation of protein modification by small protein conjugation or removal (Figure 3A). In terms of molecular function, IMRDEGs were mainly related to phosphatidyl phospholipase B activity, lysophospholipase activity, enzyme−substrate adapter activity, transcription coactivator binding, and ubiquitin ligase−substrate adapter activity (Figure 3B). Moreover, KEGG analysis showed that these genes were mainly enriched in glutathione metabolism, central carbon metabolism in cancer, and mitophagy-animal metabolism (Figure 3C).

ROC Analysis of IMRDEGs

ROC analysis for 24 IMRDEGs in the combined dataset was performed. The AUC values for 24 IMRDEGs ranged from 0.599 to 0.781 (Figure 4A-F). Among these, the top three genes with AUC value were SKP1, HNF4A, and PLA2G6. Specifically, the AUC value of SKP1 was 0.781, HNF4A was 0.746, and PLA2G6 was 0.733 (Figure 4D-F).

Construction of the Risk Score Model

To assess the diagnostic potential of the 24 IMRDEGs in PD, the logistic regression analysis was performed. The results indicated that 23 IMRDEGs were statistically significant and are displayed in the forest plot (Figure 5A). Then, the SVM algorithm was utilized to determine feature genes. We found that the SVM model achieved the smallest error and highest accuracy when utilizing the 10 genes, including HNF4A, MAP1LC3C, CHAC1, CDO1, PNPLA6, ACVR1B, GPX2, SKP1, TFR2, and ALAS2 (Figure 5B). Based on these 10 genes, the LASSO regression

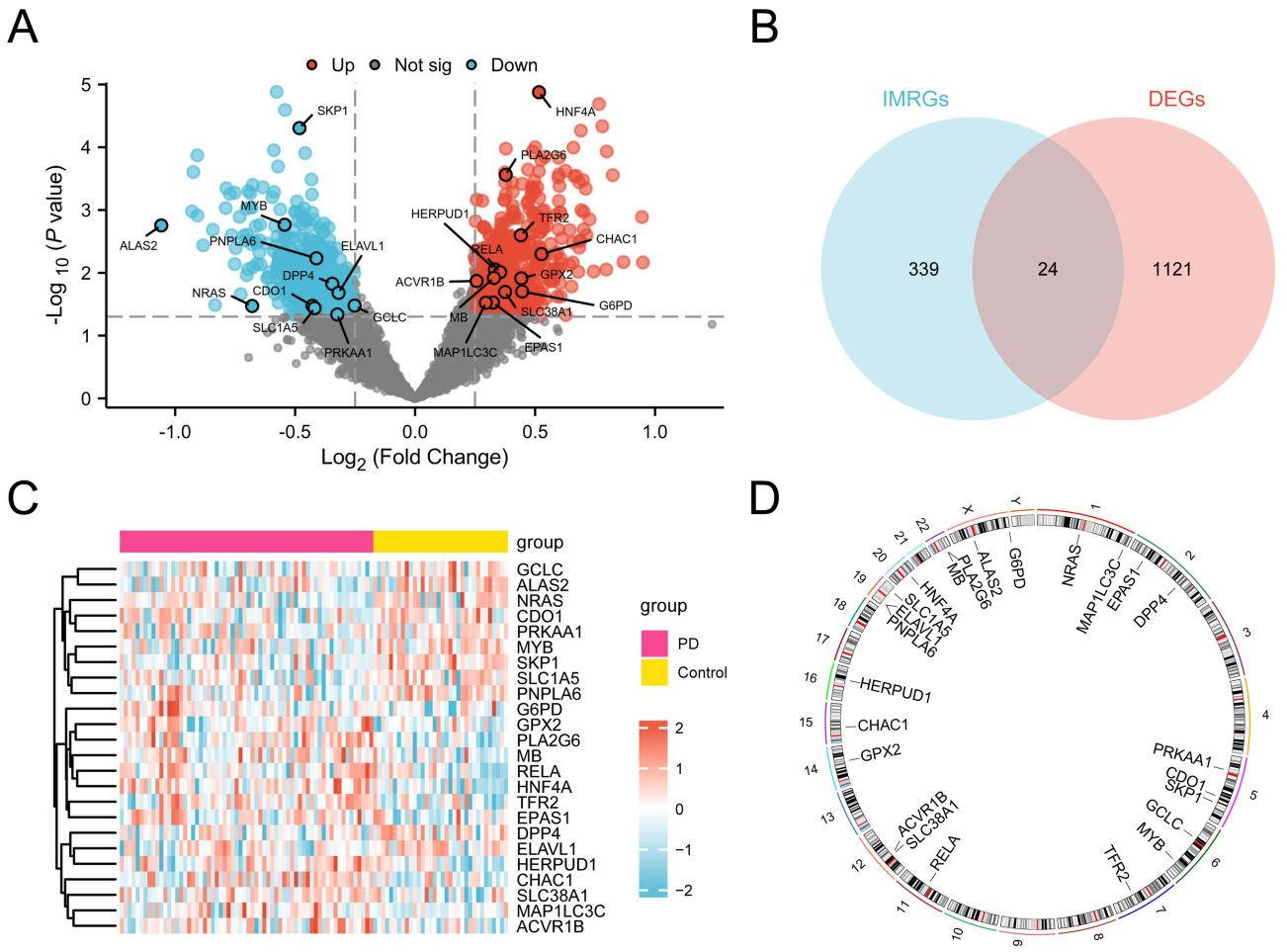

Figure 2 Identification of IMRDEGs. (**A**) Volcano plot of PD-related DEGs. (**B**) Venn diagram of PD-related DEGs and IMRGs. (**C**) Heatmap of IMRDEGs. (**D**) The locations of IMRDEGs on the chromosomes.

model was constructed (Figure 5C). The risk score model is as follows: $CHAC1 - 0.109 \times CDO1 - 0.471 \times PNPLA6 + 0.394 \times ACVR1B + 0.284 \times GPX2 - 1.130 \times SKP1 + 0.356 \times TFR2 - 0.057 \times ALAS2$. The 10 feature genes were considered as the key genes.

Evaluation of the Risk Score Model

The AUC value of the risk score model was 0.953 in the combined dataset, indicating excellent diagnostic performance in distinguishing PD samples from healthy controls (Figure 6A). Significant correlations were found between five pairs of key genes (Figure 6B), with the highest correlation observed between ALAS2 and GPX2 (correlation coefficient = 0.330), followed by TFR2 and PNPLA6 (correlation coefficient = 0.314) (Figure 6C and D). Friends analysis revealed that MAP1LC3C plays an important role in PD as the gene closest to the cut-off value (cut-off value = 0.6, Figure 6E). A nomogram for the risk score model was created based on 10 key genes (Figure 6F). The calibration curve indicated a good agreement between the predicted and observed outcomes (Figure 6G). The DCA showed a higher clinical benefit of the risk score model for PD patients (Figure 6H).

Drug Prediction

The potential drugs for 10 key genes were predicted using the CTD. A total of 49 drugs targeted to 9 key genes were predicted, and the mRNA-drug regulatory network was visualized by the Cytoscape software (Figure 7).

https://doi.org/10.2147/NDT.S511671

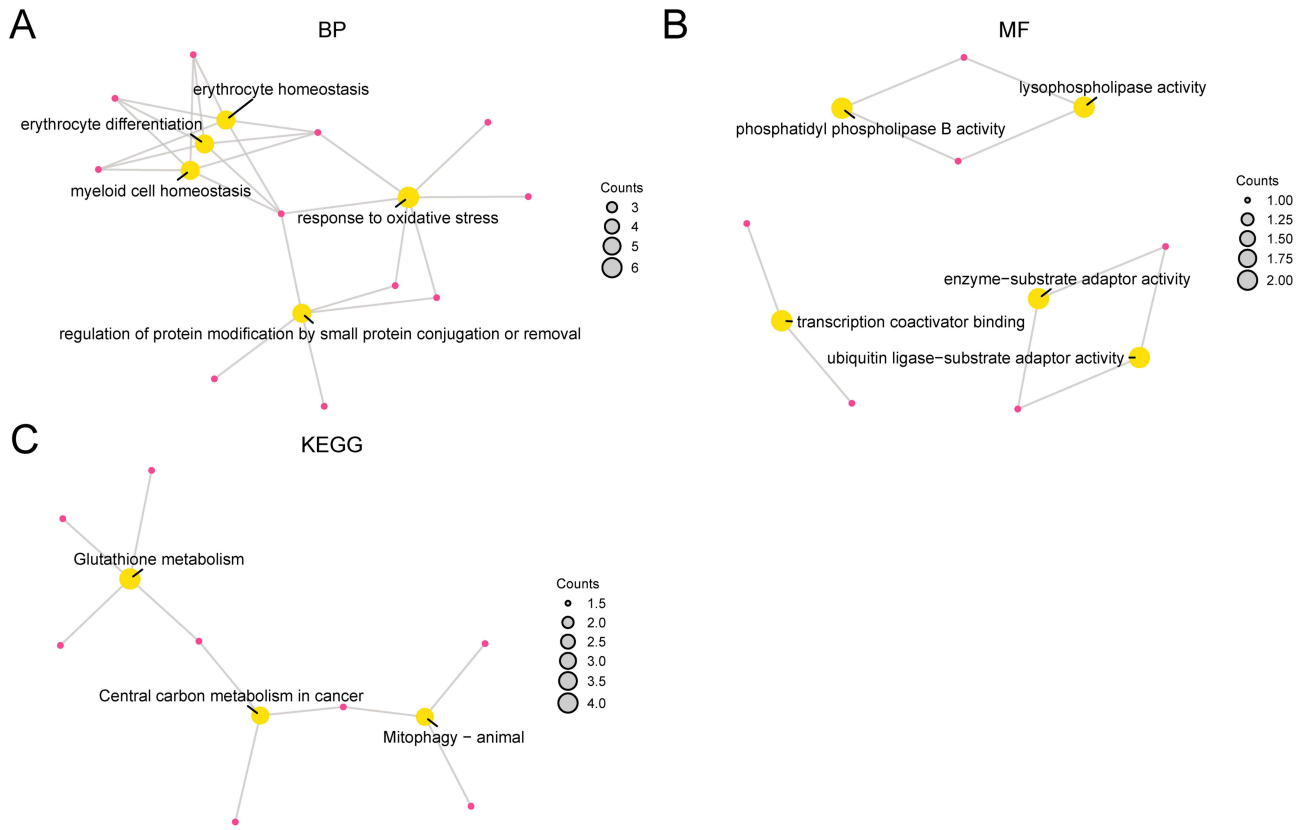

Figure 3 GO and KEGG analysis results of IMRDEGs. (**A**) GO enrichment results for biological process. (**B**) GO enrichment results for molecular function. (**C**) KEGG pathway enrichment results. Yellow nodes represent terms, red nodes represent molecules, and the connecting lines represent the relationships between terms and molecules.

RT-qPCR Validation

HNF4A, CHAC1, CDO1, PNPLA6, SKP1 and ALAS2 were randomly selected for RT-qPCR validation. All primers used for RT-qPCR detection are shown in Table S2. RT-qPCR results showed that HNF4A and CHAC1 had a tendency to be up-regulated in PD (Figure 8A and B), while CDO1, PNPLA6, SKP1 and ALAS2 had a tendency to be down-regulated in PD (Figure 8C-F). Among them, the expressions of PNPLA6 and SKP1 were significantly different between the control and PD groups. In addition, the lack of significance of most genes may be caused by small sample size and sample heterogeneity. Therefore, it is necessary to expand the sample size for further study in the later stage.

Discussion

PD is an irreversible neurodegenerative disorder characterized by the degeneration of dopaminergic neurons.[32] Its complex pathogenesis and the absence of reliable biological markers pose challenges for early diagnosis, contributing to substantial medical and societal burdens.[9] Iron dysregulation has been identified as a key mechanism underlying PD.[33] However, the specific role of IMRGs in PD pathogenesis remains to be fully understood. In this study, we identified 24 IMRDEGs in the combined dataset. Functional enrichment analysis revealed that these genes were involved in oxidative stress response, lysophospholipase activity, ubiquitin-protein transferase adaptor activity, and glutathione metabolism. ROC analysis of these genes showed that the AUC values ranged from 0.599 to 0.781. Furthermore, 10 key genes were selected using logistic regression and SVM algorithms, and a risk score model was constructed, which exhibited good diagnostic performance. Additionally, 49 potential drugs for 9 key genes were predicted.

In the risk score model, HNF4A, MAP1LC3C, CHAC1, CDO1, PNPLA6, ACVR1B, GPX2, SKP1, TFR2, and ALAS2 were identified as the key genes. HNF4A, a major regulatory gene of the nuclear receptor family, is predominantly expressed in the liver, pancreas, intestines, and kidneys.[34–37] Previous meta-analyses identified HNF4A as the

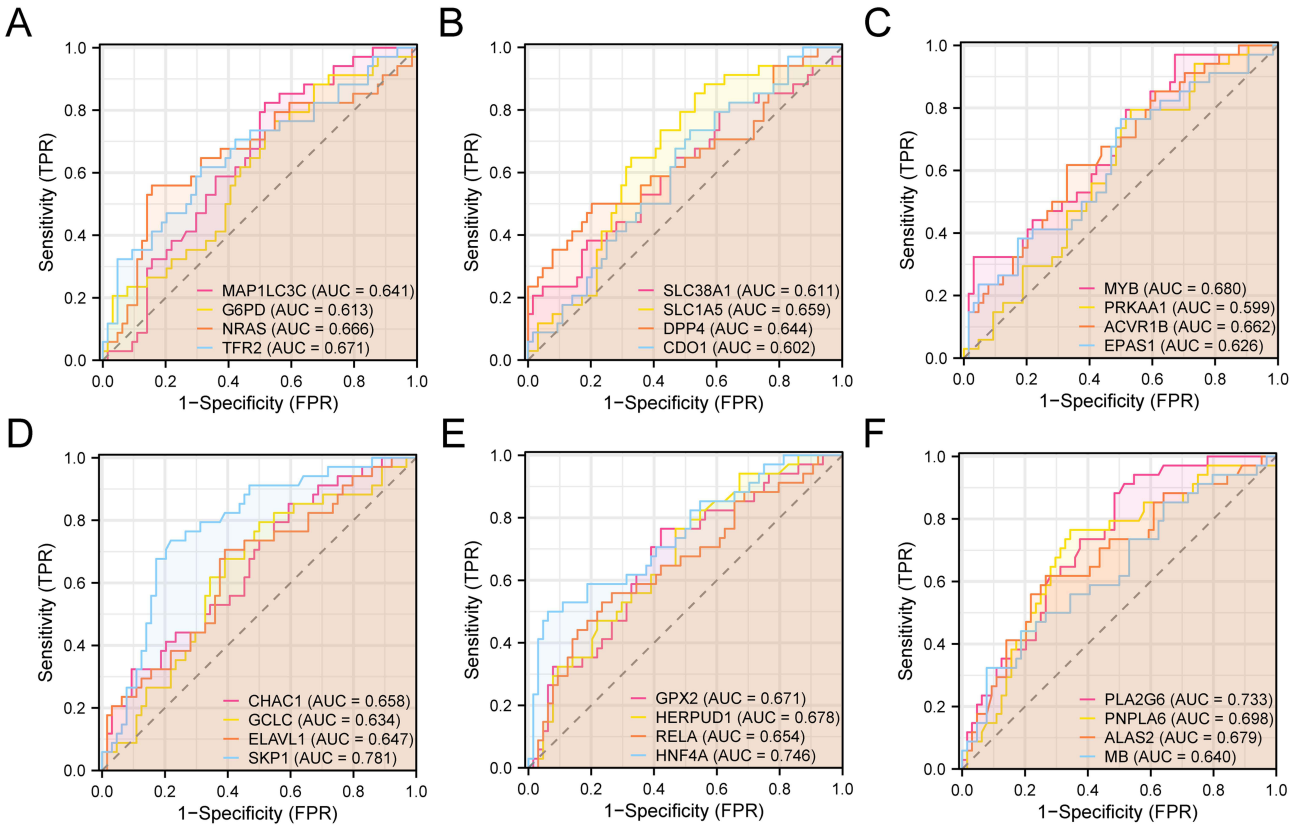

Figure 4 ROC analysis. (**A-F**) The ROC analysis of 24 IMRDEGs.

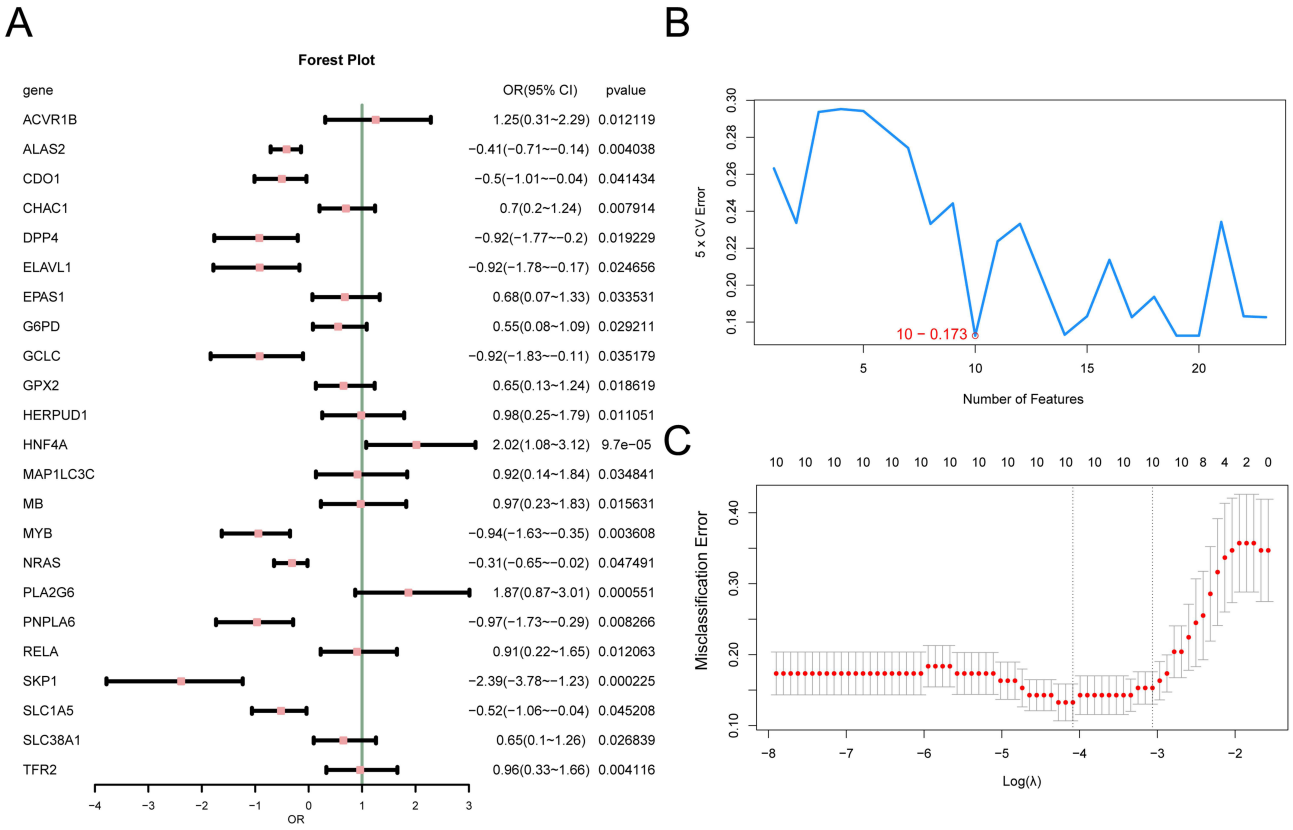

Figure 5 Construction of the risk score model. (**A**) Forest plot of the 23 IMRDEGs included in the logistic regression model of PD. (**B**) The number of genes with the minimal error rate using the SVM algorithm. (**C**) Optimal lambda values were determined using 10-fold cross-validation.

https://doi.org/10.2147/NDT.S511671

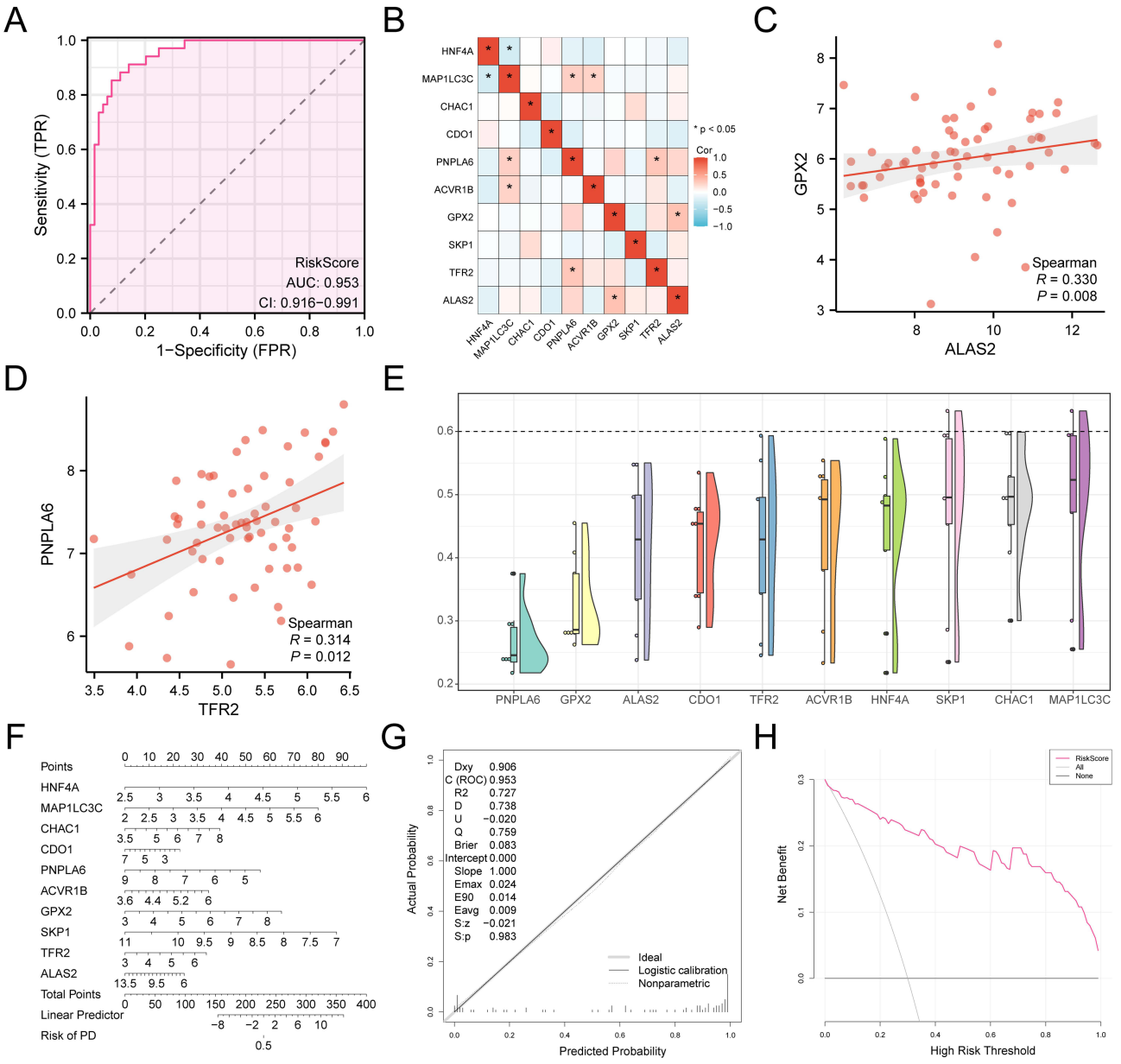

Figure 6 Evaluation of the risk score model in the combined dataset. (**A**) ROC analysis of the risk score model. (**B**) Correlation analysis among the 10 key genes. (**C**) Scatter plot of the correlation between ALAS2 and GPX2. (**D**) Scatter plot of the correlation between TFR2 and PNPLA6. (**E**) Cloud diagram of key genes. (**F**) A nomogram of the risk score model. (**G**) The calibration curve of the risk score model. (**H**) The DCA of the risk score. *, P<0.05.

most markedly upregulated gene in the blood of PD patients, which was confirmed in clinical blood samples.[38] The relative levels of HNF4A were correlated with PD severity and showed temporal variations, suggesting its potential use in tracking clinical progression.[38] PNPLA6 variants were reported to be related to PD.[39,40] SKP1 belongs to the ubiquitin E3 ligases, and the expression was decreased in PD patients.[41,42] The selective loss of dopaminergic neurons in the SN pars compacta (SNpc) is a hallmark of PD.[43] The SN contains the highest iron levels in PD patients.[44] Transferrin receptor 2 (TFR2) is a component of the mitochondrial iron transport system in PD patients.[45] In mice models with targeted deletion of TFR2 in dopaminergic neurons, TFR2 deficiency confers neuroprotection against dopaminergic degeneration and prevents iron overload associated with PD.[46] ALAS2, encoding the erythroid-specific δ-aminolevulinate synthase, initiates heme biosynthesis.[47] ALAS2-null mice exhibited substantial cytoplasmic iron accumulation in erythroblasts.[47] Meta-analysis of blood from PD patients found that the expression level of ALAS2 was reduced.[48] Moreover, the ratio of reduced to oxidized iron in the SN of PD patients was altered.[49] Glutathione (GSH),

 https://doi.org/10.2147/NDT.S511671

Figure 7 The mRNA-drug regulatory network.

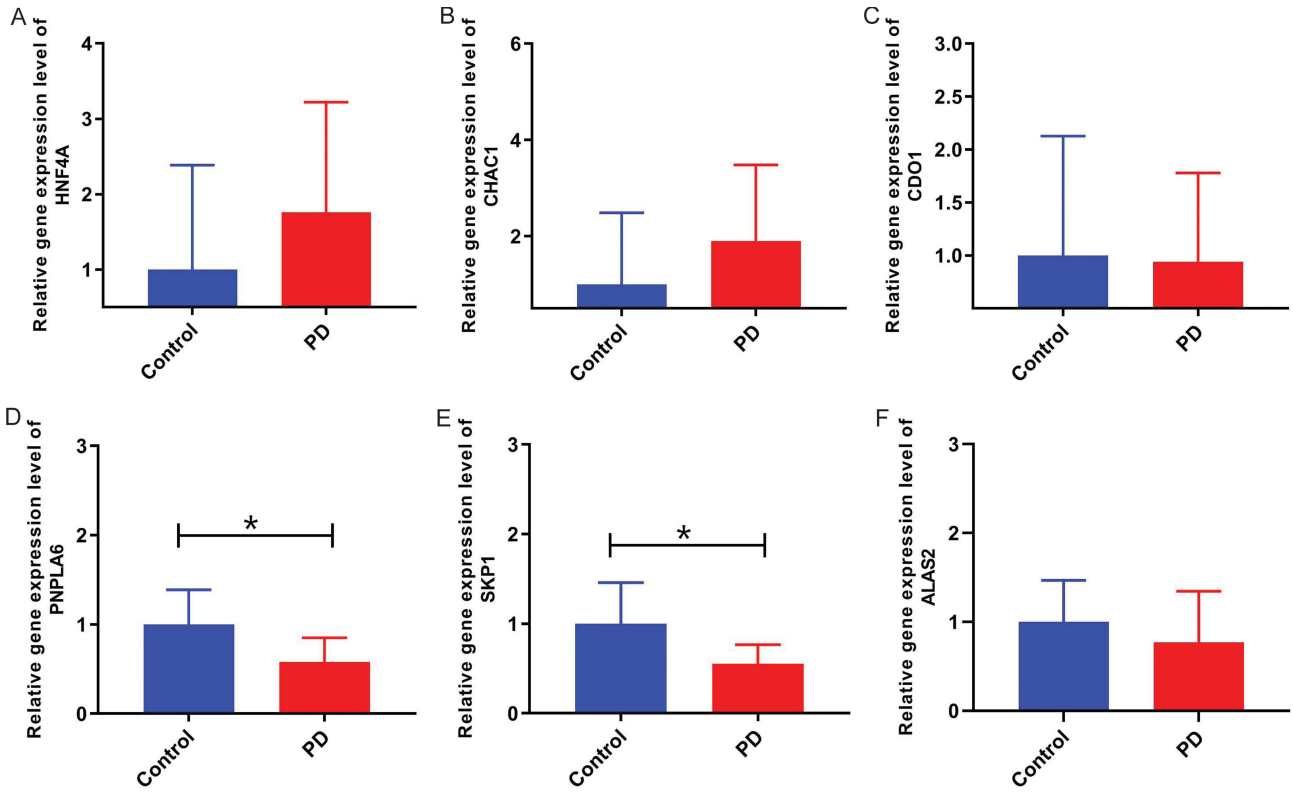

Figure 8 The expression of HNF4A (**A**), CHAC1 (**B**), CDO1 (**C**), PNPLA6 (**D**), SKP1 (**E**) and ALAS2 (**F**) was verified by RT-qPCR. * P<0.05.

https://doi.org/10.2147/NDT.S511671

a ubiquitous thiol tripeptide, protects against oxidative stress-induced damage.[50] In the PD model, overexpression of CHAC1, ChaC glutathione-specific gamma-glutamylcyclotransferase 1, led to GSH depletion, triggering oxidative stress and consequent neuronal death.[51] DJ-1 protein, encoded by a PD-associated gene, downregulates CHAC1 expression in astrocytes, reducing glutathione degradation.[52] Cysteine dioxygenase type 1 (CDO1) promotes cysteine metabolism, reducing glutathione synthesis and leading to oxidative stress.[53] In midbrain dopaminergic neuronal cells, the expression level of CDO1 was decreased as the progression of PD.[54] However, there is a lack of research on MAP1LC3C, ACVR1B and GPX2 in PD. Our study shows that MAP1LC3C, ACVR1B and GPX2 are abnormally expressed in PD, which is worthy of further exploration in the future. Overall, our findings suggest that these genes may be involved in PD progression. However, further investigations are needed to elucidate their precise roles in PD pathogenesis.

The AUC values of the 10 key genes ranged from 0.602 to 0.746, while the AUC value of the risk score model was 0.953, indicating significantly enhanced diagnostic accuracy and predictive accuracy. The pathology of PD involves multiple mechanisms, including genetic, environmental, and inflammatory.[55] Genetic risk in PD likely results from the cumulative effect of multiple common low-risk variants.[56] This multi-gene approach captures more comprehensive biological information, reflecting the PD multifactorial regulatory mechanisms. The high AUC values of the risk score risk model suggest considerable clinical potential, particularly for early screening and personalized treatment in high-risk populations.

Through mRNA-drug interaction network analysis, 49 drugs targeting 9 key genes were identified. Notably, estradiol (E2), quercetin, certain Chinese herbal components, deferoxamine (DFO), and cannabidiol (CBD) emerge as potential candidates for ameliorating the clinical symptoms of PD. The higher E2 levels were observed in the female early-onset PD, which significantly negatively correlated with the MDS Unified Parkinson's Disease Rating Scale.[57] In the rotenone-induced PD model, E2 treatment was able to modulate abnormal antioxidant enzymes and lipid peroxidation levels to normal, demonstrating a neuroprotective effect.[58]

Abnormal α-Syn aggregation in PD brains is associated with mitochondrial dysfunction and neuroinflammation.[59] Reducing reactive oxygen species (ROS) lessens α-syn toxicity and promotes the regeneration of dopaminergic neurons.[60] Quercetin, a naturally occurring flavonoid, significantly reduced mitochondrial DNA damage and oxidative stress in PD models.[61] The antioxidant properties and neuroprotective effects of traditional Chinese medicine are gaining attention in PD research.[62] The neuroprotective of Ping-wei-san plus herbal decoction and Tongtian oral liquid were reported in the PD models.[63,64]

DFO, an iron chelator, presents a promising therapeutic approach for neurodegenerative diseases.[65] Studies on PD patients treated with DFO have shown that iron overload in the SN was reduced.[66,67] CBD is a non-psychoactive compound found in cannabis plants.[68] Preclinical and preliminary clinical studies suggested that CBD has therapeutic potential for non-motor symptoms of PD, such as sleep disturbances, cognitive deficits, psychosis, depression, and anxiety.[68] Hence, the development and research of these drugs may offer novel therapeutic approaches for PD.

However, there are some limitations in this study. Firstly, our data were acquired from the public datasets, and the specific mechanism of the identified key genes is still unclear. A large number of in vivo and in vitro experiments will be conducted in the later stage to explore the specific molecular mechanisms of the key genes identified. Furthermore, risk score model based on public data lack validation in clinical samples. Therefore, a large number of clinical samples will be collected to verify the accuracy of the risk score model.

Conclusion

In summary, our study identified 24 IMRDEGs in PD patients through bioinformatics. Then, 10 key genes were identified to construct the risk score model, demonstrating a strong diagnostic value for PD. Our research may provide new insights into PD pathogenesis and contribute to the development of novel diagnostic and therapeutic strategies.

Abbreviations

PD, Parkinson's disease; IMRGs, iron metabolism-related genes; DEGs, differentially expressed genes; IMRDEGs, iron metabolism-related DEGs; SN, substantia nigra; RT-qPCR, real-time polymerase chain reaction; GEO, Gene Expression Omnibus; PCA, Principal component analysis; FC, fold change; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of

Genes and Genomes; FDR, false discovery rate; ROC, Receiver operating characteristic; AUC, area under the curve; SVM, support vector machine; LASSO, least absolute shrinkage and selection operator; DCA, decision curve analysis; CTD, Comparative Toxicogenomics Database; SNpc, SN pars compacta; TFR2, Transferrin receptor 2; CDO1, Cysteine dioxygenase type 1; DFO, deferoxamine; ROS, Reducing reactive oxygen species.

Data Sharing Statement

We searched for PD public gene expression data from GEO (https://www.ncbi.nlm.nih.gov/geo) databases. The accession number are GSE6613, GSE54536 and GSE22491.

Ethics Approval and Consent to Participate

This study was approved by the Ethics Committee of Suzhou Hospital of Anhui Medical University (2024LL028). This study complied with the Declaration of Helsinki. Written informed consent was obtained from all participants.

Funding

There is no funding to report.

Disclosure

The authors declare that they have no competing interests.

References

1. Arkinson C, Walden H. Parkin function in Parkinson's disease. *Science*. 2018;360(6386):267–268. doi:10.1126/science.aar6606
2. Michel Patrick P, Hirsch Etienne C, Hunot S. Understanding dopaminergic cell death pathways in Parkinson disease. *Neuron*. 2016;90(4):675–691. doi:10.1016/j.neuron.2016.03.038
3. Bhidayasiri R, Rattanachaisit W, Phokaewvarangkul O, Lim TT, Fernandez HH. Exploring bedside clinical features of parkinsonism: a focus on differential diagnosis. *Parkinsonism Related Disord*. 2019;59:74–81. doi:10.1016/j.parkreldis.2018.11.005
4. Li L, Wang Z, You Z, Huang J. Prevalence and influencing factors of depression in patients with Parkinson's disease. *Alpha Psychiatry*. 2023;24(6):234–238. doi:10.5152/alphapsychiatry.2023.231253
5. Menkü BE, Akın S. Diagnostic transitions from primary psychiatric disorders to underlying medical conditions: a 5-year retrospective survey from a university hospital sample. *Alpha Psychiatry*. 2024;25(2):226–232.
6. Xu T, Dong W, Liu J, et al. Disease burden of Parkinson's disease in China and its provinces from 1990 to 2021: findings from the global burden of disease study 2021. *Lancet Regional Health Western Pacific*. 2024;46:101078. doi:10.1016/j.lanwpc.2024.101078
7. Jankovic J, Tan EK. Parkinson's disease: etiopathogenesis and treatment. *J Neurol Neurosurg*. 2020;91(8):795–808. doi:10.1136/jnnp-2019-322338
8. Santos WT, Katchborian-Neto A, Viana GS, et al. Metabolomics unveils disrupted pathways in parkinson's disease: toward biomarker-based diagnosis. *ACS Chem Neurosci*. 2024;15(17):3168–3180. doi:10.1021/acschemneuro.4c00355
9. Tolosa E, Garrido A, Scholz SW, Poewe W. Challenges in the diagnosis of Parkinson's disease. *Lancet Neurol*. 2021;20(5):385–397. doi:10.1016/S1474-4422(21)00030-2
10. Schrag A, Ben-Shlomo Y, Quinn N. How valid is the clinical diagnosis of Parkinson's disease in the community? *J Neurol Neurosurg*. 2002;73(5):529–534. doi:10.1136/jnnp.73.5.529
11. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg*. 1992;55(3):181–184. doi:10.1136/jnnp.55.3.181
12. Yao Z, Jiao Q, Du X, et al. Ferroptosis in Parkinson's disease —— the iron-related degenerative disease. *Ageing Res Rev*. 2024;101:102477. doi:10.1016/j.arr.2024.102477
13. Kulaszyńska M, Kwiatkowski S, Skonieczna-żydecka K. The iron metabolism with a specific focus on the functioning of the nervous system. *Biomedicines*. 2024;12(3):595. doi:10.3390/biomedicines12030595
14. Gao G, You L, Zhang J, Chang Y-Z, Yu P. Brain iron metabolism. *Redox Balance Neurol Dis*. 2023;12(6):1289.
15. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics*. 2007;23(14):1846–1847. doi:10.1093/bioinformatics/btm254
16. Scherzer CR, Eklund AC, Morse LJ, et al. Molecular markers of early Parkinson's disease based on gene expression in blood. *Proc Natl Acad Sci USA*. 2007;104(3):955–960. doi:10.1073/pnas.0610204104
17. Alieva A, Shadrina MI, Filatova EV, et al. Involvement of endocytosis and alternative splicing in the formation of the pathological process in the early stages of Parkinson's disease. *Biomed Res Int*. 2014;2014:718732. doi:10.1155/2014/718732
18. Mutez E, Larvor L, Leprêtre F, et al. Transcriptional profile of Parkinson blood mononuclear cells with LRRK2 mutation. *Neurobiol Aging*. 2011;32(10):1839–1848. doi:10.1016/j.neurobiolaging.2009.10.016
19. Yao J, Chen X, Liu X, Li R, Zhou X, Qu Y. Characterization of a ferroptosis and iron-metabolism related lncRNA signature in lung adenocarcinoma. *Can Cell Inter*. 2021;21(1):340. doi:10.1186/s12935-021-02027-2
20. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;28(6):882–883. doi:10.1093/bioinformatics/bts034

21. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. doi:10.1093/nar/gkv007
22. Ben salem K, Ben Abdelaziz A. Principal component analysis (PCA). *La Tunisie medicale Avril*. 2021;99(4):383–389.
23. Zhang H, Meltzer P, Davis S. RCircos: an R package for Circos 2D track plots. *BMC Bioinf*. 2013;14(1):244. doi:10.1186/1471-2105-14-244
24. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics*. 2012;16(5):284–287. doi:10.1089/omi.2011.0118
25. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf*. 2011;12(1):77. doi:10.1186/1471-2105-12-77
26. Sanz H, Valim C, Vegas E, Oller JM, Reverter F. SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinf*. 2018;19(1):432. doi:10.1186/s12859-018-2451-4
27. Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin Epigenetics*. 2019;11(1):123. doi:10.1186/s13148-019-0730-1
28. Yu G, Li F, Qin Y, Bo X, Wu Y, Wang S. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics*. 2010;26(7):976–978. doi:10.1093/bioinformatics/btq064
29. Wu J, Zhang H, Li L, et al. A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: a population-based analysis. *Cancer Communications*. 2020;40(7):301–312. doi:10.1002/cac2.12067
30. Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and interpreting decision curve analysis: a guide for investigators. *Europ Urol*. 2018;74 (6):796–804. doi:10.1016/j.eururo.2018.08.038
31. Grondin CJ, Davis AP, Wiegers JA, et al. Predicting molecular mechanisms, pathways, and health outcomes induced by Juul e-cigarette aerosol chemicals using the comparative toxicogenomics database. *Curr Res Toxicol*. 2021;2:272–281. doi:10.1016/j.crtox.2021.08.001
32. Jiang Y, Qi Z, Zhu H, et al. Role of the globus pallidus in motor and non-motor symptoms of Parkinson's disease. *Neural Regeneration Res*. 2025;20(6):1628–1643. doi:10.4103/NRR.NRR-D-23-01660
33. Xiao Z, Wang X, Pan X, Xie J, Xu H. Mitochondrial iron dyshomeostasis and its potential as a therapeutic target for Parkinson's disease. *Exp Neurol*. 2024;372:114614. doi:10.1016/j.expneurol.2023.114614
34. Hunter AL, Poolman TM, Kim D, et al. HNF4A modulates glucocorticoid action in the liver. *Cell Rep*. 2022;39(3):110697. doi:10.1016/j.celrep.2022.110697
35. Marable SS, Chung E, Adam M, Potter SS, Park JS. Hnf4a deletion in the mouse kidney phenocopies Fanconi renotubular syndrome. *JCI Insight*. 2018;3(14). doi:10.1172/jci.insight.97497
36. Brunton H, Caligiuri G, Cunningham R, et al. HNF4A and GATA6 loss reveals therapeutically actionable subtypes in pancreatic cancer. *Cell Rep*. 2020;31(6):107625. doi:10.1016/j.celrep.2020.107625
37. Gu W, Huang X, Singh PNP, et al. A MTA2-SATB2 chromatin complex restrains colonic plasticity toward small intestine by retaining HNF4A at colonic chromatin. *Nat Commun*. 2024;15(1):3595. doi:10.1038/s41467-024-47738-y
38. Santiago JA, Potashkin JA. Network-based metaanalysis identifies HNF4A and PTBP1 as longitudinally dynamic biomarkers for Parkinson's disease. *Proceedings Nat Acad Sci United States of America*. 2015;112(7):2257–2262. doi:10.1073/pnas.1423573112
39. Sen K, Finau M, Ghosh P. Bi-allelic variants in PNPLA6 possibly associated with Parkinsonian features in addition to spastic paraplegia phenotype. *J Neurol*. 2020;267(9):2749–2753. doi:10.1007/s00415-020-10028-w
40. Kazanci S, Witt J, Su K, et al. PNPLA6-related disorder with levodopa-responsive parkinsonism. *Movement Disord Clin Pract*. 2023;10 (2):338–340. doi:10.1002/mdc3.13632
41. Dabool L, Hakim-Mishnaevski K, Juravlev L, Flint-Brodsly N, Mandel S, Kurant E. Drosophila Skp1 homologue Skpa plays a neuroprotective role in adult brain. *iScience*. 2020;23(8):101375. doi:10.1016/j.isci.2020.101375
42. Salas-Leal AC, Salas-Pacheco SM, Hernández-Cosaín EI, et al. Differential expression of PSMC4, SKP1, and HSPA8 in Parkinson's disease: insights from a Mexican mestizo population. *Front Mol Neurosci*. 2023;16. 10.3389/fnmol.2023.1298560
43. Guatteo E, Berretta N, Monda V, Ledonne A, Mercuri NB. Pathophysiological features of nigral dopaminergic neurons in animal models of Parkinson's disease. *Int J Mol Sci*. 2022;23(9):4508. doi:10.3390/ijms23094508
44. Homayoon N, Pirpamer L, Franthal S, et al. Nigral iron deposition in common tremor disorders. *Movement Disord*. 2019;34(1):129–132. doi:10.1002/mds.27549
45. Mastroberardino PG, Hoffman EK, Horowitz MP, et al. A novel transferrin/TfR2-mediated mitochondrial iron transport system is disrupted in Parkinson's disease. *Neurobiol Dis*. 2009;34(3):417–431. doi:10.1016/j.nbd.2009.02.009
46. Milanese C, Gabriels S, Barnhoorn S, et al. Gender biased neuroprotective effect of Transferrin Receptor 2 deletion in multiple models of Parkinson's disease. *Cell Death Differ*. 2021;28(5):1720–1732. doi:10.1038/s41418-020-00698-4
47. Harigae H, Nakajima O, Suwabe N, et al. Aberrant iron accumulation and oxidized status of erythroid-specific δ-aminolevulinate synthase (ALAS2)–deficient definitive erythroblasts. *Blood*. 2003;101(3):1188–1193. doi:10.1182/blood-2002-01-0309
48. Santiago JA, Potashkin JA. Blood transcriptomic meta-analysis identifies dysregulation of hemoglobin and iron metabolism in Parkinson. *Disease*. 2017;9.
49. Faucheux BA, Martin ME, Beaumont C, et al. Lack of up-regulation of ferritin is associated with sustained iron regulatory protein-1 binding activity in the substantia nigra of patients with Parkinson's disease. *J Neurochemistr*. 2002;83(2):320–330. doi:10.1046/j.1471-4159.2002.01118.x
50. Smeyne M, Smeyne RJ. Glutathione metabolism and Parkinson's disease. *Free Radic Biol Med*. 2013;62:13–25. doi:10.1016/j.freeradbiomed.2013.05.001
51. Do JH. Neurotoxin-induced pathway perturbation in human neuroblastoma SH-EP Cells. *Mol Cells*. 2014;37(9):672–684. doi:10.14348/molcells.2014.0173
52. Ge Y, Zheng X, Mao S, Zhang Q, Hu G, Wei Y. DJ-1 inhibits glutathione degradation by downregulating CHAC1 expression in astrocytes. *Neuroscience Res*. 2022;184:62–69. doi:10.1016/j.neures.2022.08.006
53. Chen M, Zhu JY, Mu WJ, Guo L. Cysteine dioxygenase type 1 (CDO1): its functional role in physiological and pathophysiological processes. *Genes Dis*. 2023;10(3):877–890. doi:10.1016/j.gendis.2021.12.023
54. Wang J, Duhart HM, Xu Z, Patterson TA, Newport GD, Ali SF. Comparison of the time courses of selective gene expression and dopaminergic depletion induced by MPP+ in MN9D cells. *Neurochem Int*. 2008;52(6):1037–1043. doi:10.1016/j.neuint.2007.10.017

55. Le Heron C, MacAskill M, Mason D, et al. A multi-step model of parkinson's disease pathogenesis. *Movement Disord*. 2021;36(11):2530–2538. doi:10.1002/mds.28719
56. Dehestani M, Liu H, Gasser T. Polygenic risk scores contribute to personalized medicine of Parkinson's disease. *J Personalized Med*. 2021;11 (10):1030. doi:10.3390/jpm11101030
57. Bovenzi R, Conti M, Simonetta C, et al. Contribution of testosterone and estradiol in sexual dimorphism of early-onset Parkinson's disease. *J Neural Transm*. 2024.
58. Shen D, Tian X, Zhang B, Song R. Mechanistic evaluation of neuroprotective effect of estradiol on rotenone and 6-OHDA induced Parkinson's disease. *Pharmacol Rep*. 2017;69(6):1178–1185. doi:10.1016/j.pharep.2017.06.008
59. Das SS, Jha NK, Jha SK, Verma PRP, Ashraf GM, Singh SK. Neuroprotective role of quercetin against alpha-synuclein-associated hallmarks in Parkinson's disease. *Curr Neuropharmacol*. 2023;21(7):1464–1466. doi:10.2174/1570159X21666221221092250
60. Muhammad F, Liu Y, Zhou Y, Yang H, Li H. Antioxidative role of traditional Chinese medicine in Parkinson's disease. *J Ethnopharmacol*. 2022;285:114821. doi:10.1016/j.jep.2021.114821
61. Wang WW, Han R, He HJ, et al. Administration of quercetin improves mitochondria quality control and protects the neurons in 6-OHDA-lesioned Parkinson's disease models. *Aging*. 2021;13(8):11738–11751. doi:10.18632/aging.202868
62. Chen P, Zhang J, Wang C, et al. The pathogenesis and treatment mechanism of Parkinson's disease from the perspective of traditional Chinese medicine. *Phytomedicine*. 2022;100:154044. doi:10.1016/j.phymed.2022.154044
63. Dongjie S, Rajendran RS, Xia Q, et al. Neuroprotective effects of Tongtian oral liquid, a Traditional Chinese Medicine in the Parkinson's disease-induced zebrafish model. *Biomed Pharmacother*. 2022;148:112706. doi:10.1016/j.biopha.2022.112706
64. Li D, H-j Y, G-j H, R-y Y, A-m X, X-y L. Mechanisms of the Ping-wei-san plus herbal decoction against Parkinson's disease: multiomics analyses. *Front Nutrition*. 2023;9:945356.
65. You L, Wang J, Liu T, et al. Targeted brain delivery of rabies virus glycoprotein 29-modified deferoxamine-loaded nanoparticles reverses functional deficits in parkinsonian mice. *ACS nano*. 2018;12(5):4123–4139. doi:10.1021/acsnano.7b08172
66. Devos D, Moreau C, Devedjian JC, et al. Targeting chelatable iron as a therapeutic modality in Parkinson's disease. *Antioxid Redox Signaling*. 2014;21(2):195–210. doi:10.1089/ars.2013.5593
67. Grolez G, Moreau C, Sablonnière B, et al. Ceruloplasmin activity and iron chelation treatment of patients with Parkinson's disease. *BMC Neurol*. 2015;15(1):74. doi:10.1186/s12883-015-0331-3
68. Crippa JAS, Hallak JEC, Zuardi AW, Guimarães FS, Tumas V, Dos Santos RG. Is cannabidiol the ideal drug to treat non-motor Parkinson's disease symptoms? *European Archiv Psychiatr Clin Neurosci*. 2019;269(1):121–133. doi:10.1007/s00406-019-00982-6