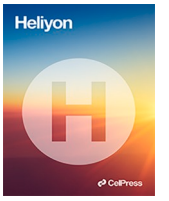
Heliyon 11 (2025) e41831

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Glycolytic pathways: The hidden regulators in Parkinson's disease

Jing Shen [a,*], Ensheng Yao [a], Weidong Tian [b], Jia He [c], Yukai Gu [c], Dong Zhao [b,**]

[a] Department of Neurology, the First Affiliated Hospital of Shihezi University, China
[b] Department of Neurosurgery, the First Affiliated Hospital of Shihezi University, China
[c] Department of Preventive Medicine, School of Medical, Shihezi University, China

ARTICLE INFO

Keywords:
Parkinson's disease
Glycolysis
Hub genes

ABSTRACT

Parkinson's disease (PD) is a widespread neurodegenerative condition [1]; however, its association with glycolysis, specifically the activity of genes related to glycolysis, has not yet been explored. We downloaded 3 datasets related to PD from the GEO database and identified the glycolytic genes related to PD. Subsequently, GO and KEGG enrichment analyses were conducted. We constructed a PD diagnosis model using the SVM algorithm for differentially expressed glycolysis-related genes and verified the model with LASSO regression analysis. Next, we constructed a regulatory network of genes that were differentially expressed with respect to glycolysis. Finally, the amount of immune cell infiltration was analyzed in PD samples, and the correlation between differential genes and immune cells was calculated. A total of 64 differentially expressed glycolytic genes associated with PD were screened. Then, a GO analysis was conducted, followed by KEGG and GASE enrichment analyses. Within the established PD diagnostic model, 26 genes that were differentially expressed and linked to glycolysis showed strong statistical significance. After further screening, a diagnostic model for PD including seven key genes was established. Further analysis showed that ABHD5 most strongly correlated with neutrophils ($r = 0.507$). The key gene SMAD3 was strongly negatively associated with gamma delta T cells ($r = -0.488$). This research offered a theoretical foundation for the association between glycolysis and PD. Seven glycolytic genes were identified as significantly linked to PD and warrant additional research.

1. Introduction

Parkinson's disease (PD) is characterized by slow movement, muscle stiffness, and tremors [1], which negatively affect patients' quality of life. Illness development is influenced by various factors that usually affect individuals in their middle to later years, including genetic factors, environmental influences, and neurodegenerative changes [2]. Currently, drug and surgical approaches are the primary treatment options for PD, but they have significant limitations and only attenuate symptoms. Consequently, understanding the pathogenesis of PD and pinpointing therapeutic targets is vital.

Glycolysis is an important metabolic pathway in cells that produces energy by decomposing glucose, which can be used by cells to carry out various life activities, such as synthesizing proteins and releasing energy [3]. This process usually occurs in the cytoplasm and requires multiple enzymes. Under normal circumstances, the ATP produced by glycolysis is used to maintain normal cell function.

* Corresponding author.
** Corresponding author.
E-mail addresses: sub0621marine@163.com (J. Shen), shzzhaodong112@163.com (D. Zhao).

https://doi.org/10.1016/j.heliyon.2025.e41831
Received 9 April 2024; Received in revised form 7 January 2025; Accepted 8 January 2025
Available online 17 January 2025

However, the disruption of glycolysis can cause energy deficiency, acidosis, and even cell death.

According to previous studies, the brains of PD patients have protein aggregations of alpha-synuclein (α-Syn), which might interfere with signal transmission between neurons, resulting in the occurrence of PD [4,5]. Some studies have suggested that α-Syn aggregation may indirectly affect glycolytic pathways through pathways such as mitochondrial function, oxidative stress, and inflammatory responses [6]. Furthermore, several studies have indicated glycolytic-related alterations in the brain tissue of individuals with PD that affect processes such as mitochondrial dysfunction and oxidative stress [7–9]. Glycolysis may promote neuronal death and the loss of dopaminergic neurons by affecting the energy metabolism and antioxidant capacity of neurons. Additionally, glycolysis may also affect the function of neurons by affecting the synthesis and release of neurotransmitters. These changes might lead to functional impairment of neurons and increase the deposition of α-Syn in the brains of patients, thereby affecting motor control and cognitive function [8]. Recent advancements in research have shed light on the intricate relationship between glycolysis and PD, underscoring the metabolic dysfunction that may underpin the disease's pathophysiology. Naeem et al. [10] highlighted glycolysis as a potentially crucial factor in PD, suggesting that disruptions in glycolytic pathways can lead to inadequate ATP production, exacerbating oxidative stress and promoting neurodegeneration. The authors emphasize that understanding these metabolic changes could lead to innovative therapeutic approaches that target glycolytic processes, thereby providing neuroprotection and possibly altering the progression of PD. Specifically, they discuss the importance of enhancing glycolytic capacity as a means to support energy metabolism in dopamine-producing neurons, which are notably vulnerable in Parkinson's patients. In addition, some studies have explored the application of glycolysis in the treatment of PD [11]. For example, it is possible to reduce the symptoms in patients with PD by improving mitochondrial function and reducing oxidative stress [12]. While there is no definitive proof of a direct cause-and-effect connection between glycolysis and PD, it has been suggested that abnormalities in glycolysis may contribute to PD development and progression [13]. Hence, further investigation is required to delve into the connection between glycolysis and PD for a better comprehension of disease onset and the development of improved therapies.

This study examined the correlation between glycolysis and the onset and prognosis of PD using biological information, constructed and validated a diagnostic model of PD, explored biomarkers with diagnostic and prognostic value, and provided new ideas for the selection of subsequent treatment targets.

2. Materials and methods

2.1. Data download

The PD-related datasets GSE100054 [14] GSE6613 [15,16] GSE54536 [17] were obtained from the GEO database using the R package GEOquer [18]. These datasets included samples from *Homo sapiens* specifically sourced from blood tissue. The chip platforms of GSE100054, GSE6613, GSE54536 were GPL23126, GPL96, and GSE10558, respectively, and detailed information is provided in Table 1. This study included both PD and control samples. The dataset GSE100054 contained ten PD samples and nine control samples; dataset GSE6613 contained 50 PD and 22 control samples; and dataset GSE54536 contained four PD and four control samples. None of the three datasets involved in our study (GSE100054, GSE6613, and GSE54536) contained stage information for Parkinson's disease, nor did they address drug treatment and comorbidities. Data for this study are available in a public, open access repository. The datasets can be found in the GEO (https://www.ncbi.nlm.nih.gov/).

Glycolysis related genes (GRGs) were collected using the GeneCards Database and Molecular Signatures Database (MSigDB) [19,20]. GeneCards provides comprehensive genetic information about humans. In the search, we used the term "Glycolysis" as the keyword and only retained the "Protein Coding" GRGs that were found, and a total of 3062 GRGs were obtained. Similarly, we searched the MSigDB database with the word "Glycolysis" as the keyword and retrieved the BIOCARTA GLYCOLYSIS PATHWAY, BIOCARTA FEEDER PATHWAY, BIOCARTA ETC PATHWAY, and BIOCARTA KREB PATHWAY gene sets containing a total of 51 GRGs. In total, 3063 genes associated with glycolysis were identified after merging and removing duplicates.

Combined datasets were obtained using R packages sva and GSE54536. The combined datasets comprised 64 PD and 35 control samples [21]. Combined geospatial datasets were standardized, probes were annotated, and normalization was carried out with the R package limma [8,22].

Table 1
Information about GEO microarray Chips.

	GSE100054	GSE6613	GSE54536
Platform	GPL23126	GPL96	GPL10558
Species	Homo Sapiens	Homo Sapiens	Homo Sapiens
Tissue	Blood	Blood	Blood
Samples in PD group	10	50	4
Samples in Normal group	9	22	4
Reference	PMID: 29223072	PMID: 18669654 PMID: 17215369	PMID: 24804238

GEO, Gene Expression Omnibus.

2.2. Glycolysis genes differentially expressed in Parkinsonism

Using the limma R package, we analyzed the differences between genes in the PD and control groups in the combined GEO datasets. The cutoff for recognizing differentially expressed genes (DEGs) was determined at 0.30 < value <| logFC | and p <| 0.05. Among them, differentially expressed genes meeting the criteria 0.30 < value <| logFC | and p <| 0.05 were considered upregulated genes. Genes with logFC < −0.30 and p value < 0.05 were classified as downregulated genes.

By analyzing the GEO dataset using variance analysis, we identified glycolysis-related differentially expressed genes (GRDEGs) associated with PD by selecting genes with |logFC | > 0.3 and p < 0.05. The intersection of DEGs related to glycolysis was mapped using Venn diagrams to identify GRDEGs. The R package ggplot2 was used to visualize the results of the differential analysis, creating a volcano map and a differential ranking map. Additionally, the R package heatmap was used to generate a heatmap displaying the expression values of the GRDEGs.

The Kyoto Encyclopedia of Genes and Genomes (KEGG)-based analysis of enriched gene ontology (GO) terms and pathways.

Gene ontology (GO) [23] terms are used for large-scale functional enrichment studies and are categorized as biological processes (BP), molecular functions (MF), and cell components (CC) [24]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) [25] contains information on genomes, diseases, and drugs. The cluster Profiler package was used to analyze the GO and KEGG enrichment of GRDEGs [26]. Statistically significant items were identified based on screening criteria with a p value < 0.05 and a false discovery rate (FDR) value (q value) < 0.25.

2.3. The process of gene set enrichment analysis (GSEA)

Genes from the combined GEO datasets were initially categorized as PD and control based on their phenotypic significance using GSEA [27]. GSEA was conducted on all genes in the combined GEO datasets using the cluster Profiler R package [28]. The GSEA parameters included a seed of 2020 with gene sets containing a minimum of 10 genes and a maximum of 500 genes. The screening criteria for GSEA in MSigDB included adjusted p < 0.05 and FDR value (q value) < 0.25, with p-values corrected using the Benjamini-Hochberg (BH) method [20].

We categorized genes in the PD samples into high- and low-risk groups based on the median least absolute shrinkage and selection operator (LASSO) risk score and then conducted differential analysis using the R package limma. The R package ggplot2 was used to visualize the distribution of volcanic plots for the PD sample group, highlighting genes with |logFC| > 0.3 and p < 0.05 in both high- and low-risk categories. Subsequently, a heatmap of the gene expression values was generated using the same criteria for the R package heatmap. Finally, the R package cluster Profiler for PD samples of all the genes in the GSEA was used with parameters set as follows: seed for 2020, each gene set with a minimum of 10, contains the number of genes containing the number of genes for a maximum of 500. Through MSigDB Database access to c2 gene sets. Cp. All. V2022.1. Hs. Symbols, The GMT GSEA, utilized screening criteria for GSEA that included a p value < 0.05 and an FDR value (q value) < 0.25.

2.4. Construction of PD diagnostic model

Logistic regression was used to analyze the GRDEGs to generate PD diagnostic models from the combined GEO datasets. The dependent variable, which was a binary variable representing the PD group versus the control group, was analyzed using logistic regression. A significance level of 0.01 was utilized to identify GRDEGs and develop a logistic regression model. A forest plot was constructed to illustrate the GRDEGs incorporated within the logistic regression model, based on their molecular expression profiles.

Subsequently, a support vector machine (SVM) model [24] using the identified GRDEGs was developed [29]. We then conducted a LASSO regression analysis using the R package glmnet [25] on GRDEGs that were part of the SVM model. Through linear regression analysis, The LASSO regression method addresses the issue of overfitting through the inclusion of a penalty term, represented by the absolute value of the slope coefficient (lambda), leading to enhanced generalization. The analysis of LASSO regression involved diagnostic models and variable trajectory plots. The GRDEGs were included in the diagnostic model.

2.5. Validation of PD diagnostic model

To illustrate the interconnections between important genes, a nomogram was developed based on the outcomes of a logistic regression study [26]. A calibration curve was derived from the LASSO regression analysis to assess the accuracy and differentiation of the MD diagnostic model. Decision curve analysis (DCA) maps were created using the R package ggDCA with key genes from the combined GEO datasets [30]. Diagnostic tests, molecular markers, and clinical prediction models were evaluated using the DCA. Subsequently, the pROC R package was used to generate receiver operating characteristic (ROC) curves and to compute the area under the curve (AUC) in the combined GEO datasets. Based on the following formula, the LASSO risk score was computed to evaluate the diagnostic impact of LASSO expression on the occurrence of PD:

$$\text{risk}Score = \sum_i Coefficient\ (gene_i) * mRNA\ Expression\ (gene_i)$$

Furthermore, LASSO risk scores were used to categorize PD samples into high- and low-risk categories. A comparison map was generated to further explore the differences in gene expression between the high- and low-risk categories of PD samples, based on the levels of the key genes. Expression levels were assessed using the ROC curve of the key genes, and the area under the curve (AUC) was

calculated using the R package pROC. Typically, the AUC was within the range of 0.5 1. An AUC value > 0.5 suggests a the molecule's expression tends to facilitate the event, with a higher AUC indicating a more effective diagnostic outcome. 0.5–0.7 AUC values indicated low accuracy, 0.7–0.9 AUC values indicated moderate accuracy, and AUC values > 0.9 indicated high accuracy.

2.6. Construction of regulatory network

MicroRNAs (miRNAs) play important roles in biological development and evolution [31]. Several miRNAs can control a single target gene, and conversely, a single target gene can be influenced by multiple miRNAs. The regulatory networks of key genes and miRNAs were analyzed by retrieving the miRNAs associated with key genes from the TarBase database [32]. The regulatory network of mRNA and miRNA was depicted using Cytoscape software [33].

After transcription, transcription factors (TFs) interact with their target mRNAs to regulate gene expression. TFs [34] retrieved from ChIPBase [35] and htftarget database were pooled to analyze the regulatory effects of TFs on the key genes. Cytoscape software facilitated the visualization of the mRNA-TF regulatory network [36].

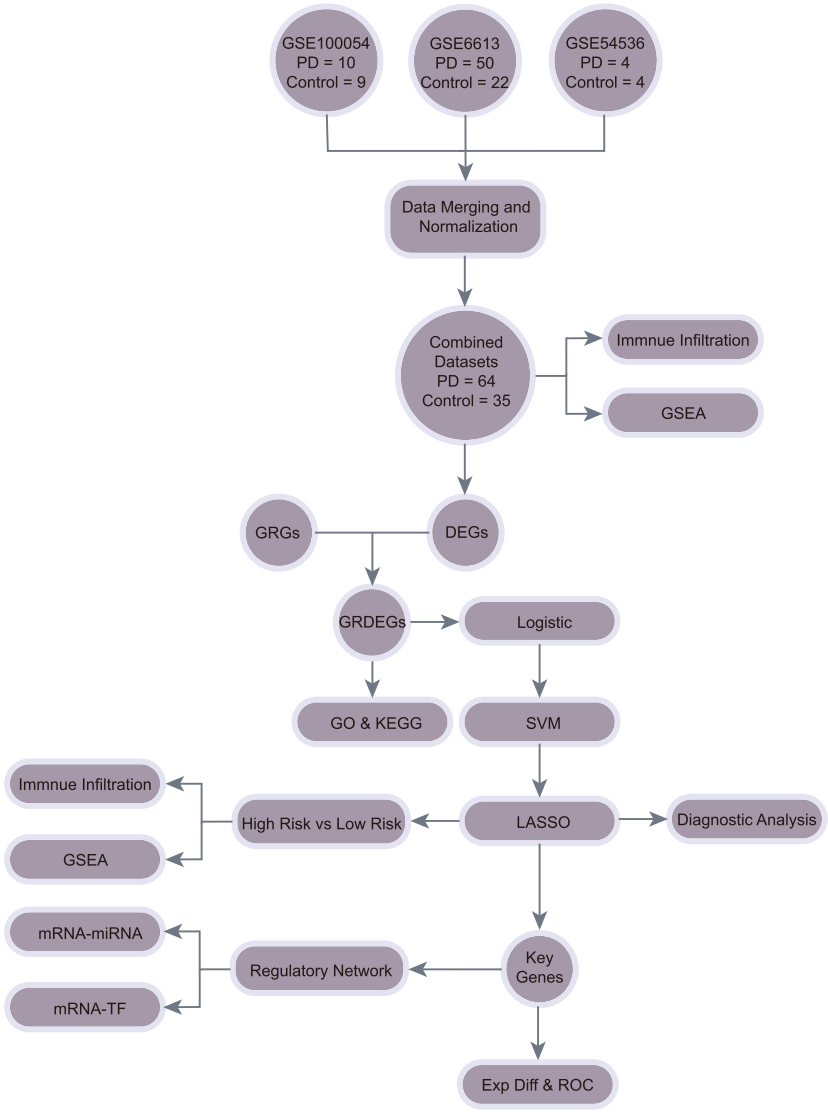


Fig. 1. A flow chart is presented for the comprehensive analysis of GRDEGs.
PD, Parkinson's disease; GSEA, gene set enrichment analysis; DEGs, differentially expressed genes; GRGs, glycolysis related genes; GRDEGs, glycolysis related differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; SVM, support vector machine; LASSO, least absolute shrinkage and selection operator; miRNA, micro ribonucleic acid; mRNA, messenger ribonucleic acid; ROC curve, receiver operating characteristic curve; TF, transcription factor; Exp Diff, expression difference analysis.

2.7. Immune infiltration analysis of the integrated dataset

An immune cell mixture can be predicted using CIBERSORT based on a linear support vector regression analysis of the transcriptome expression matrix [37] (http://cibersort.stanford.edu/). Using the LM22 gene matrix, the CIBERSORT algorithm [38] was employed to compute the matrix of immune cell infiltration from the merged GEO datasets. Infiltration matrix outcomes were acquired from data without immune cell enrichment scores above zero. The pheatmap R package was used to generate correlation heat maps displaying the correlation analysis findings for CIBERSORT immune cells and key genes in immune cells.

Simultaneously, we used the CIBERSORT algorithm in conjunction with the LM22 gene matrix to assess immune cell infiltration levels in PD samples. Specific immune cell infiltration matrix results were obtained by excluding data with immune cell enrichment scores above zero. The correlation analysis results of CIBERSORT immune cells and key genes in immune cells were shown with correlation heat maps created in R using the pheatmap package.

The proportion of individual immune cell infiltrates in a single sample was assessed using single-sample gene set enrichment analysis (ssGSEA). In published articles on tumor immune infiltration, the algorithm obtained 28 gene sets related to various types of immune cells present in tumors [39]. Different types of human immune cells have been identified based on their gene expression profiles, including activated B cells, activated CD4 T cells, and activated CD8 T cells. Enrichment scores from the ssGSEA analysis obtained with the R package GSVA were used to indicate the level of infiltration of individual immune cell types in each sample within the combined GEO datasets. Correlation analysis findings for ssGSEA immune cells and key genes were displayed using correlation heat maps created using the R package pheatmap.

Additionally, we computed the enrichment scores of the MD samples using ssGSEA in the R package GSVA to indicate the level of immune cell infiltration in each sample. Visualization of the correlation analysis findings for ssGSEA immune cells and key genes in immune cells was performed using the heatmap function in the R package. When the correlation coefficients (r values) were below 0.3, the correlation was considered weak; between 0.3 and 0.5, it was moderate; and between 0.5 and 0.8, it was strong.

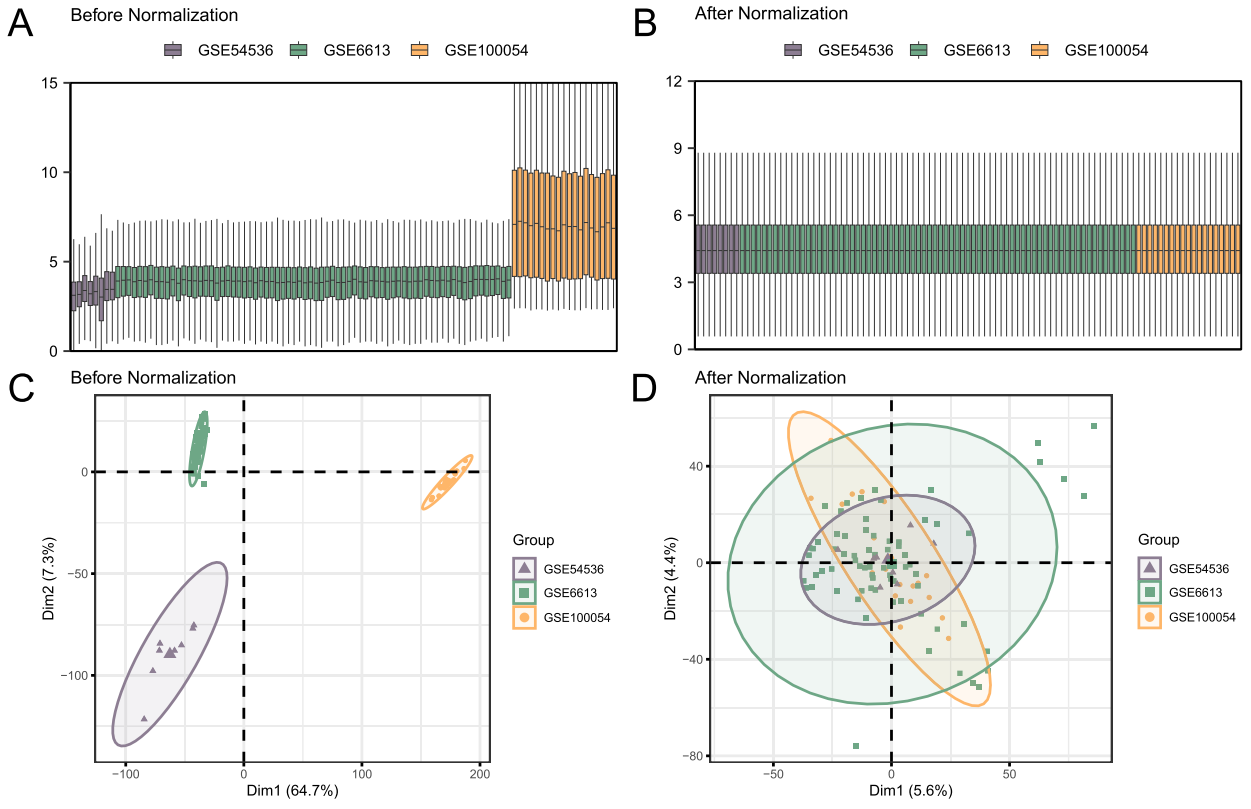


Fig. 2. Elimination of batch effects in GSE54536, GSE6613, and GSE100054.
A. Box plot of the combined GEO dataset distribution before batch processing. B. Post-batch integrated GEO datasets (combined datasets) distribution boxplots. C. PCA plot of the datasets before debatching. D. PCA map of the combined GEO datasets after batch processing. PD, Parkinson's disease; PCA, principal component analysis. The PD dataset GSE54536 is shown in purple, the PD dataset GSE6613 is shown in green, and the PD dataset GSE100054 is shown in orange.

2.8. Statistical analysis

R (version 4.2.2) was used to process and analyze the data. Continuous variables are presented as mean ± SD. The Wilcoxon rank-sum test was used to compare two groups. Spearman's correlation coefficients were calculated for various molecules, with significance determined by a P-value <0.05 (see Fig. 1).

3. Results

3.1. Technology roadmap

3.1.1. Merging of Parkinson's data sets

The process began with the use of the R package sva to address the batch effect in the datasets of PD GSE100054, GSE6613, and GSE54536, culminating in the development of the combined GEO dataset. Boxplots and PCA plots were used to compare the datasets pre and post batch effect removal (Fig. 2A–D). The removal of the batch effect was confirmed by the distribution boxplot and PCA plot results, indicating its effective elimination from the PD dataset samples [40].

3.1.2. Analysis of genes that are expressed differently in relation to PD and glycolysis

Among the unified GEO datasets, the MD and control groups were separated into the PD and control groups using the R package

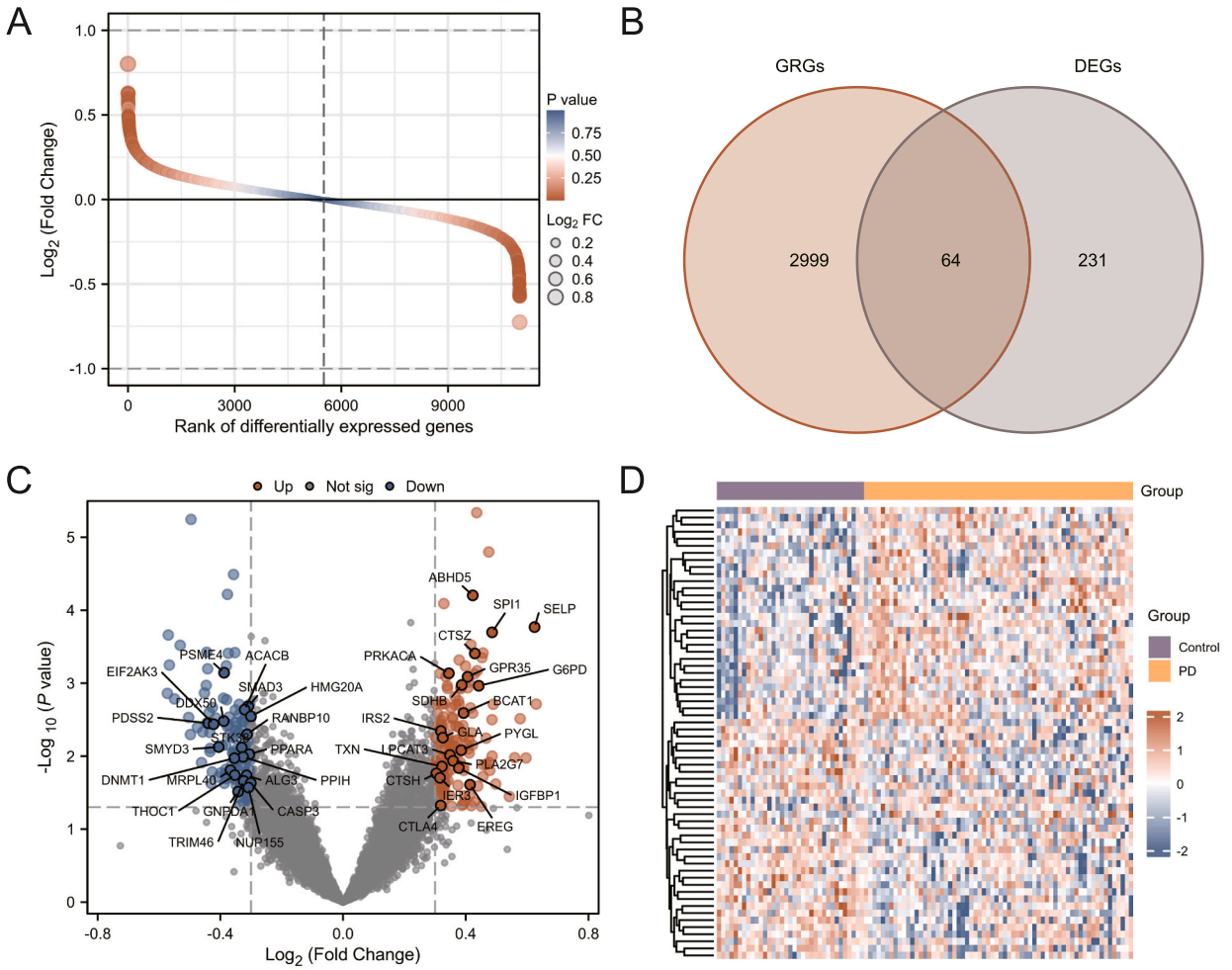


Fig. 3. Differential gene expression analysis for combined datasets.
A. Comparison of gene expression differences between the PD group and the control group in combined GEO datasets using a ranking map. B. Venn diagram showing the overlap between DEGs and GRGs in the combined GEO datasets. C. Volcano plot comparing gene expression differences between the PD and control using the integrated GEO datasets. D. Heat map displaying the expression levels of genes related to glycolysis that are differentially expressed in the combined GEO datasets. PD, Parkinsons disease; DEGs, differentially expressed genes; GRGs, glycolysis related genes; GRDEGs, glycolysis related differentially expressed genes. Purple is the control group, and orange is the PD group.

limma [22]. A total of 295 DEGs met the criteria of | logFC | > 0.30 and p < 0.05. Among these, 157 genes were upregulated (logFC >0.30 and p < 0.05) and 138 were downregulated (logFC <0.30 and p < 0.05). By analyzing this dataset, differences in ranking and volcano diagrams (Fig. 3A and C) were generated. In glycolysis, GRDEGs must have | logFC | > 0.30 and p < 0.05. The intersection of DEGs and GRDEGs resulted in 64 differentially expressed genes associated with GRDEGs, as shown in the Venn diagram (Fig. 3B) and Table 1: SDHB, PYGL, G6PD, PRKACA, NUP155, GNPDA1, PPARA, IRS2, SMAD3, ABHD5, GPR35, MRPL40, CTLA4, STK38, PPIH, RANBP10, PSME4, HMG20A, SMYD3, ALG3, TRIM46, DDX50, TXN, EREG, IGFBP1, CASP3, ACACB, SELP, LPCAT3, GLA, BCAT1, PDSS2, DNMT1, SPI1, IER3, THOC1, EIF2AK3, PLA2G7, CTSH, CTSZ, PECAM1, CEBPD, CASP8AP2, UCHL1, MYB, EMD, LRPPRC, BMP6, IMPDH2, TGFBI, TUBA8, CBR3, NR1D2, MRPL13, GNAQ, PDE3A, LCN2, SIRPA, S100A9, ITIH3, BAG2, ARIH2, IFI27, and MYL6B. Based on the intersection results, various samples within the combined GEO datasets were categorized, and the variances in the expression of GRDEGs between the PD and control groups were examined. A heat map was generated using the R package to visualize the findings (Fig. 3D).

3.1.3. Performing enrichment analysis of GRDEGs using GO and KEGG

The connection between the 64 GRDEGs and PD was further investigated using GO and KEGG Pathway Analysis to explore the relationships among biological processes (BP), molecular functions (MF), cell components (CC), and biological pathways (Pathways). Table 2 displays the results of the GO and KEGG enrichment analyses of the 64 GRDEGs. The findings indicated that GRDEGs in PD were primarily concentrated in the regulation of protein export from the nucleus, positive regulation of proteolysis, activation of cysteine-type endopeptidase activity involved in apoptosis, and regulation of lipid localization. Biological processes such as nuclear export and protein processing (BP); secretory granule lumen, cytoplasmic vesicle lumen, vesicle lumen, nuclear inner membrane, ficolin-1-rich granule, and nuclear membrane components (CC); peptidase activator activity, phosphatase binding, glucose binding, cysteine-type endopeptidase activity, cysteine-type endopeptidase activator activity involved in apoptosis, monosaccharide binding, and other molecular functions (MF). Furthermore, it enriched the glucagon signaling pathway, PD, apoptosis, and other biological pathways. A bubble diagram (Fig. 4A) was used to visualize the outcomes of GO and KEGG enrichment analyses.

Network graphs of BP, CC, MF, and pathways were generated based on the findings of GO and KEGG pathway enrichment analyses (Fig. 4B–E). Larger nodes indicate more molecules in the entries, and the lines show relationships between the entries.

3.1.4. An overview of gene set enrichment analysis (GSEA)

GSEA was used to analyze the impact of gene expression levels in the combined GEO datasets on PD, focusing on the biological processes associated with the expression of all genes in the datasets. The relationship between the affected cellular elements and molecular activities performed (Fig. 5A) is detailed in Table 3. The findings indicated a significant enrichment of all genes in the combined GEO datasets in various biological functions and signaling pathways, including MAP3K8 Tpl2 Dependent MAPK1 3 Activation (Fig. 5B), FCERI-mediated NF-κB Activation (Fig. 5C), MAPK6 MAPK4 Signaling (Fig. 5D), Regulation of TP53 Activity Through Phosphorylation (Fig. 5E), IL4 2pathway (Fig. 5F), II8 CXCR1 Pathway (Fig. 5G), and others.

3.1.5. Construction of PD diagnostic model

In the initial study, logistic regression was performed using the 64 GRDEGs to evaluate their diagnostic significance for PD. The

Table 2
Results of GO and KEGG enrichment analysis for GRDEGs.

ONTOLOGY	ID	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0046825	4/63	31/18800	3.37 e−06	0.003771	0.002861
BP	GO:0045862	9/63	369/18800	3.65 e−06	0.003771	0.002861
BP	GO:0006919	5/63	77/18800	5.9 e−06	0.004069	0.003086
BP	GO:1905952	6/63	155/18800	1.31 e−05	0.005641	0.004279
BP	GO:0051168	6/63	156/18800	1.36 e−05	0.005641	0.004279
BP	GO:0016485	7/63	243/18800	1.65 e−05	0.005696	0.004321
CC	GO:0034774	8/64	322/19594	9.7 e−06	0.000597	0.00051
CC	GO:0060205	8/64	325/19594	1.04 e−05	0.000597	0.00051
CC	GO:0031983	8/64	327/19594	1.08 e−05	0.000597	0.00051
CC	GO:0005637	4/64	61/19594	4.7 e−05	0.001937	0.001656
CC	GO:0101002	5/64	185/19594	0.000345	0.011392	0.009739
CC	GO:0031965	6/64	300/19594	0.000436	0.01199	0.01025
MF	GO:0016504	3/64	43/18410	0.000448	0.052435	0.041396
MF	GO:0019902	5/64	193/18410	0.000555	0.052435	0.041396
MF	GO:0005536	2/64	11/18410	0.000641	0.052435	0.041396
MF	GO:0004197	4/64	120/18410	0.000807	0.052435	0.041396
MF	GO:0008656	2/64	17/18410	0.001564	0.061724	0.04873
MF	GO:0048029	3/64	71/18410	0.001935	0.061724	0.04873
KEGG	hsa04922	5/49	107/8164	0.000424	0.082863	0.076598
KEGG	hsa05012	7/49	266/8164	0.000958	0.082863	0.076598
KEGG	hsa04210	5/49	136/8164	0.001262	0.082863	0.076598

GO, Gene Ontology; BP, Biological Process; CC, Cellular Component; MF, Molecular Function; KEGG, Kyoto Encyclopedia of Genes and Genomes; GRDEGs, Glycolysis Related Differentially Expressed Genes.

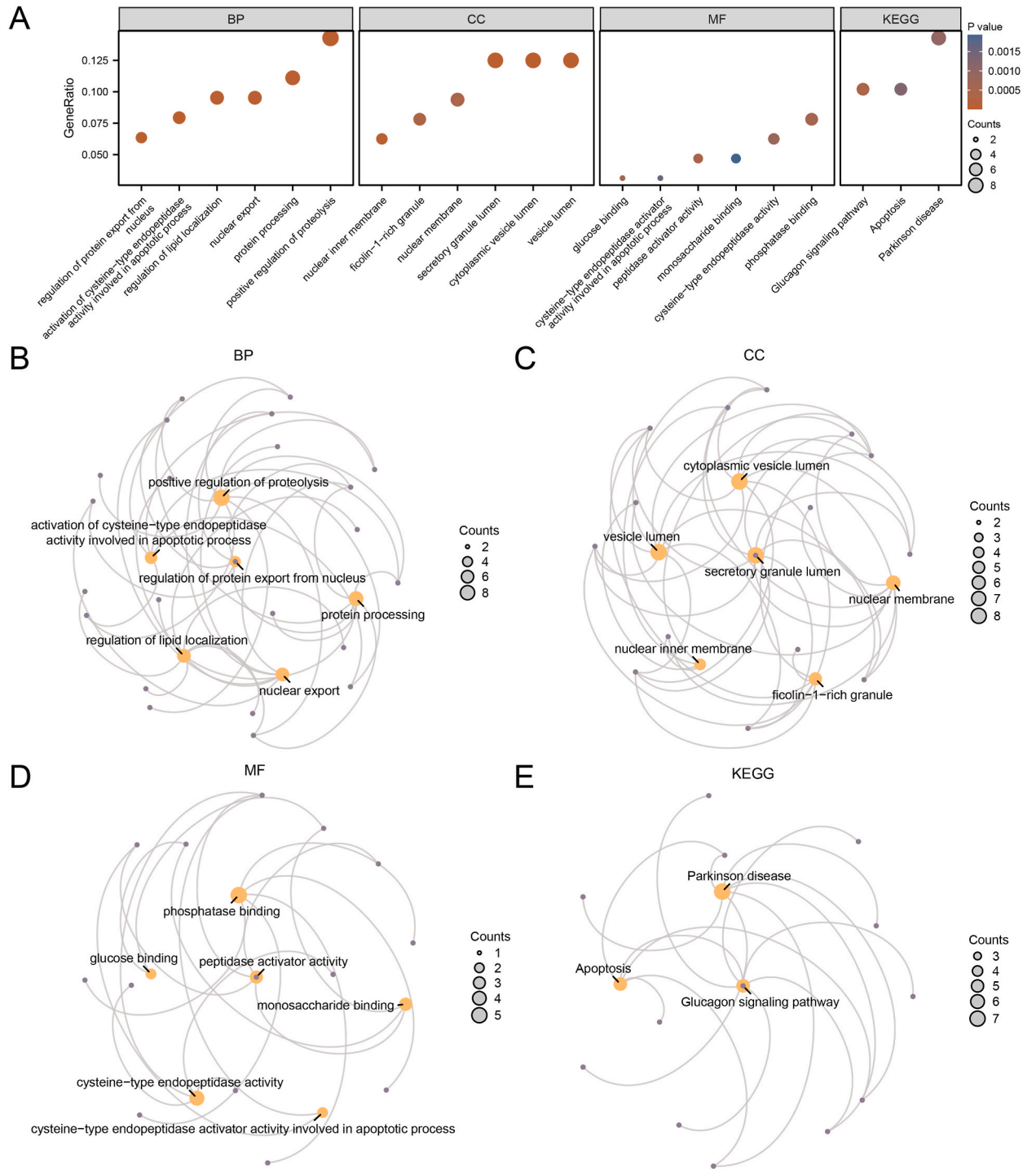


Fig. 4. Shows GRDEs enriched by GO and KEGG for growth regulation.
A. Visualization of GO and KEGG enrichment analysis results of GRDEGs: BP, CC, MF, and biological pathway (Pathway). The abscissa displays GO terms and KEGG terms. B-E. GO and KEGG enrichment analysis results of GRDEGs network diagram showing: BP (B), CC (C), MF (D), and KEGG (E). PD, Parkinsons disease; GRDEGs, glycolysis related differentially expressed genes; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, biological process; CC, cellular component; MF, molecular function. Yellow nodes represent BP, CC, MF, and KEGG entries, purple nodes represent molecules, and lines represent relationships between entries and molecules. Statistically significant screening criteria for GO and KEGG enrichment analysis included a p value < 0.05 and an FDR value (q value) < 0.25.

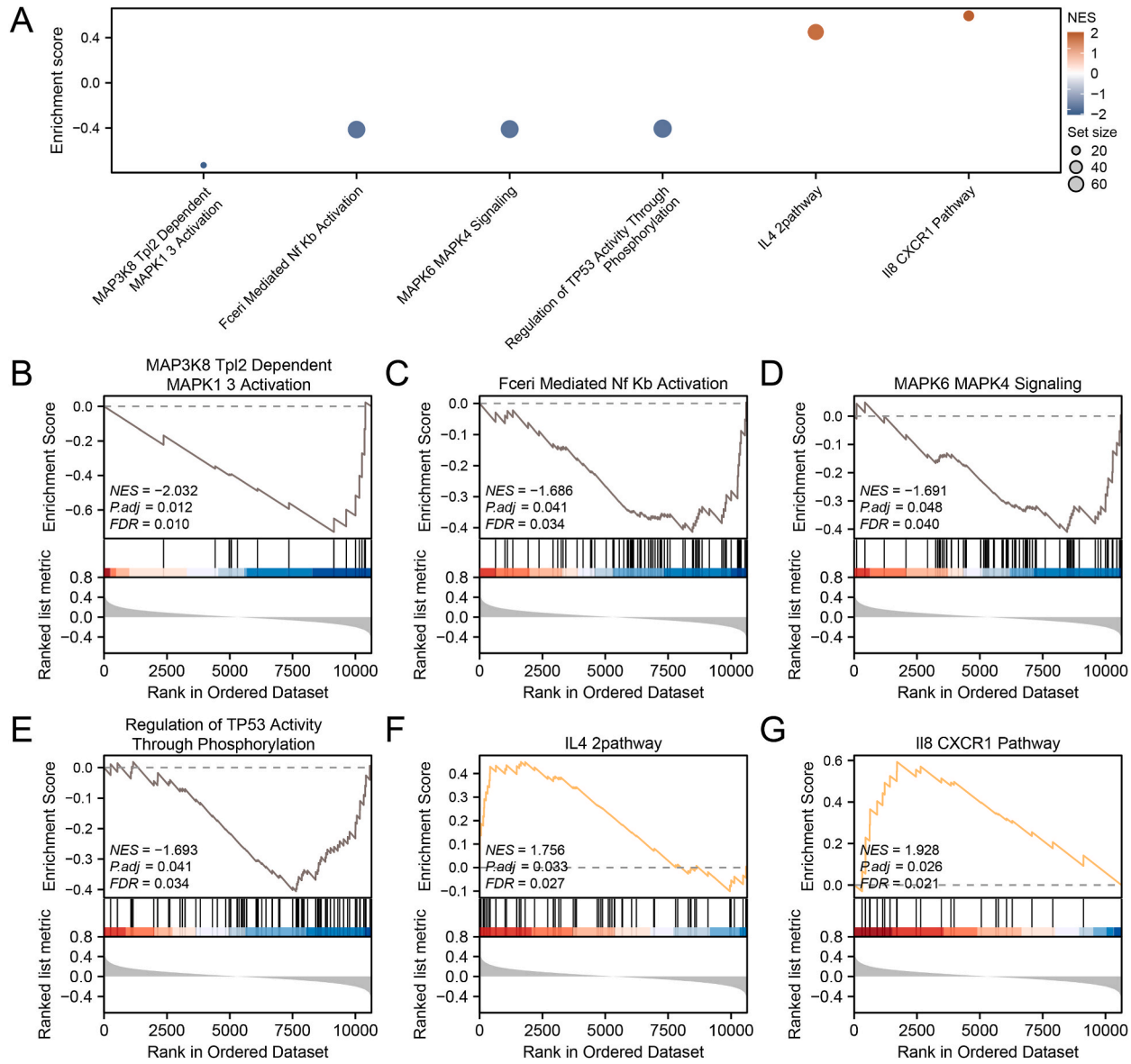


Fig. 5. A GSEA for combining datasets.

A. GSEA presents a biological function mountain map integrating the GEO datasets. B-H. GSEA revealed a significant enrichment of PD in MAP3K8 Tpl2 Dependent MAPK1 3 activation (B), FCERI-mediated NF-κB activation (C), and MAP3K8 TPL2 dependent MAPK1 3 activation (B). MAPK6 MAPK4 Signaling (D), Regulation of TP53 Activity Through Phosphorylation (E), the IL4 2pathway (F), and the Il8 CXCR1 Pathway (G). PD, Parkinson's disease; GSEA, gene set enrichment analysis. The screening criteria of gene set enrichment analysis (GSEA) were adjusted to $p < 0.05$ and FDR value (q value) < 0.25. p-value correction was performed using the Benjamini-Hochberg method (BH).

resulting logistic regression model was created and displayed using a forest plot (Fig. 6A). The findings indicated that 26 genes related to glycolysis, known as GRDEGs, were extremely statistically significant in the logistic regression model with a p value < 0.01, which were: ABHD5, ACACB, ARIH2, BCAT1, CASP8AP2, CEBPD, CTSZ, DDX50, EIF2AK3, G6PD, GLA, GNAQ, GPR35, HMG20A, IRS2, LRPPRC, MYB, PDSS2, PRKACA, PSME4, RANBP10, SDHB, SELP, SIRPA, SMAD3, and SPI1. Based on the 26 GRDEGs and SVM algorithm, we created an SVM model that identified genes with the lowest error rate (Fig. 6B) and highest accuracy (Fig. 6C). The results showed that the SVM model had the highest accuracy when the number of genes was seven, and the seven GRDEGs were LRPPRC, ABHD5, SMAD3, PSME4, GPR35, MYB, and SDHB. Subsequently, the LASSO regression model was employed to develop the diagnostic model for PD using the GRDEGs present in the 7 SVM models. Visualization included generating the LASSO regression model plot (Fig. 6D) and the LASSO variable trajectory plot (Fig. 6E). Seven glycolysis-related GRDEGs (key genes) were incorporated into the LASSO regression model: LRPPRC, ABHD5, SMAD3, PSME4, GPR35, MYB, and SDHB.

Table 3
Results of GSEA for combined datasets.

ID	Set Size	Enrichment Score	NES	pvalue	p.adjust	qvalue
REACTOME_FORMATION_OF_FIBRIN_CLOT_CLOTTING_CASCADE	38	0.643398	2.315689	1.7 e−06	0.000251	0.000206
REACTOME_PLATELET_ACTIVATION_SIGNALING_AND_AGGREGATION	227	0.460664	2.240585	1E-10	5.91 e−08	4.86 e−08
WP_MICROGLIA_PATHOGEN_PHAGOCYTOSIS_PATHWAY	37	0.62463	2.236086	5.66 e−06	0.000669	0.00055
WP_COMPLEMENT_AND_COAGULATION_CASCADES	55	0.572604	2.220592	7.69 e−07	0.00013	0.000107
REACTOME_RESPONSE_TO_ELEVATED_PLATELET_CYTOSOLIC_CA2	111	0.497844	2.209849	4.21 e−08	1.5 e−05	1.23 e−05
REACTOME_NEUTROPHIL_DEGRANULATION	379	0.429424	2.20457	1E-10	5.91 e−08	4.86 e−08
WP_COMPLEMENT_SYSTEM	89	0.512217	2.164392	6.13 e−07	0.00013	0.000107
REACTOME_INTRINSIC_PATHWAY_OF_FIBRIN_CLOT_FORMATION	22	0.667313	2.108927	0.000105	0.005632	0.004631
REACTOME_RHO_GTPASES_ACTIVATE_NADPH_OXIDASES	22	0.665865	2.104353	0.000117	0.005849	0.00481
WP_VITAMIN_K_METABOLISM_AND_ACTIVATION_OF_DEPENDENT_PROTEINS	10	0.827008	2.102748	0.000175	0.007394	0.00608
KEGG_PANTOTHENATE_AND_COA_BIOSYNTHESIS	13	0.76574	2.064589	0.000208	0.00846	0.006957
WP_MONOAMINE_TRANSPORT	28	0.608022	2.050174	0.000368	0.012053	0.009912
REACTOME_ROS_AND_RNS_PRODUCTION_IN_PHAGOCYTES	29	0.601681	2.038758	0.000348	0.012053	0.009912
PID_INTEGRIN2_PATHWAY	26	0.614933	2.033391	0.000325	0.011514	0.009468
PID_IL8_CXCR1_PATHWAY	24	0.593144	1.928192	0.001323	0.025846	0.021254
PID_IL4_2PATHWAY	57	0.449679	1.755772	0.001839	0.032674	0.026869
REACTOME_FCERI_MEDIATED_NF_KB_ACTIVATION	70	0.41376	1.68593	0.002619	0.041155	0.033843
REACTOME_MAPK6_MAPK4_SIGNALING	74	0.41011	1.69118	0.003469	0.048216	0.039649
REACTOME_REGULATION_OF_TP53_ACTIVITY_THROUGH_PHOSPHORYLATION	78	0.40556	1.69268	0.002603	0.041155	0.033843
REACTOME_MAP3K8_TPL2_DEPENDENT_MAPK1_3_ACTIVATION	15	0.72781	2.03221	0.000362	0.012053	0.009912

GSEA, Gene Set Enrichment Analysis.

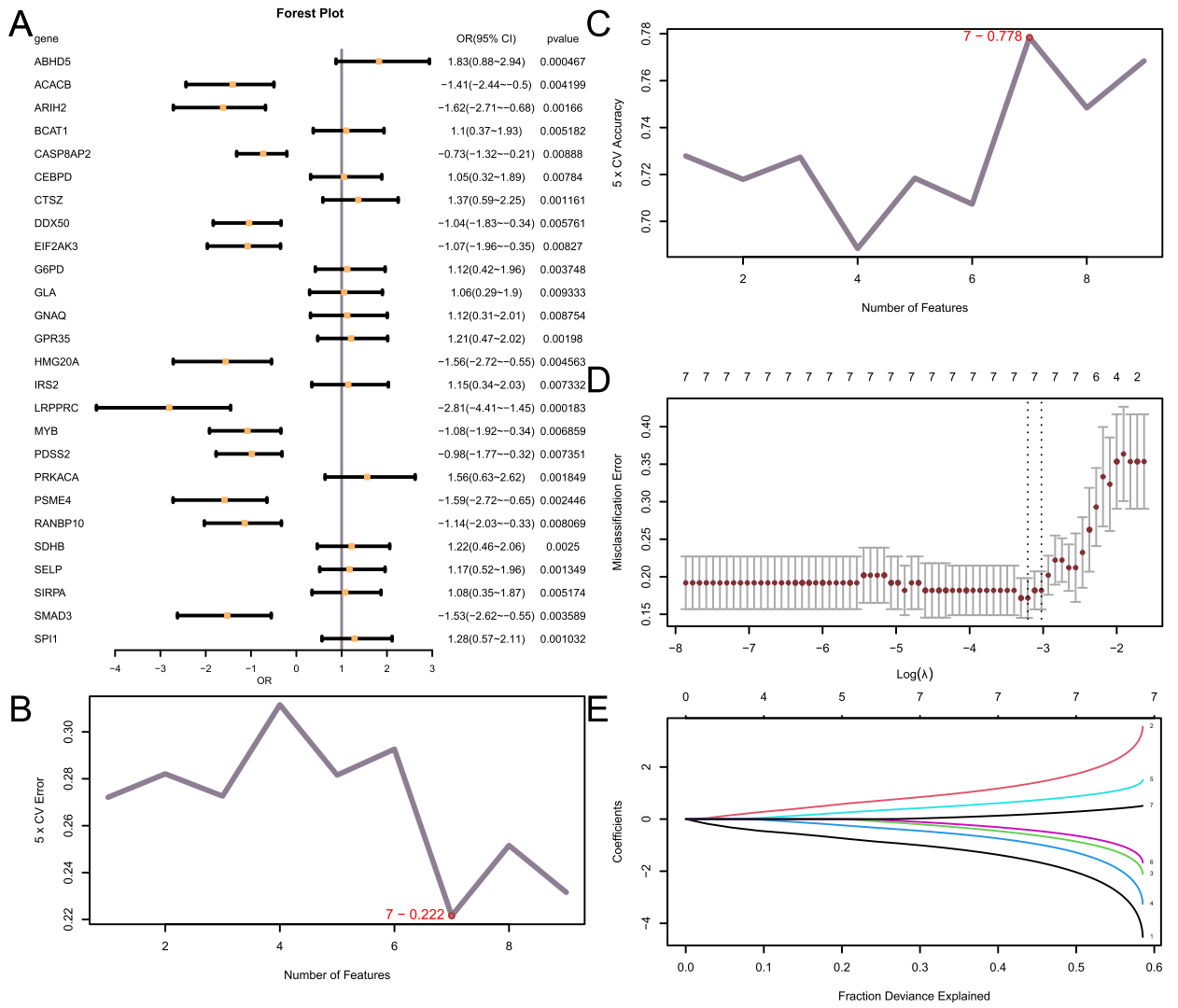


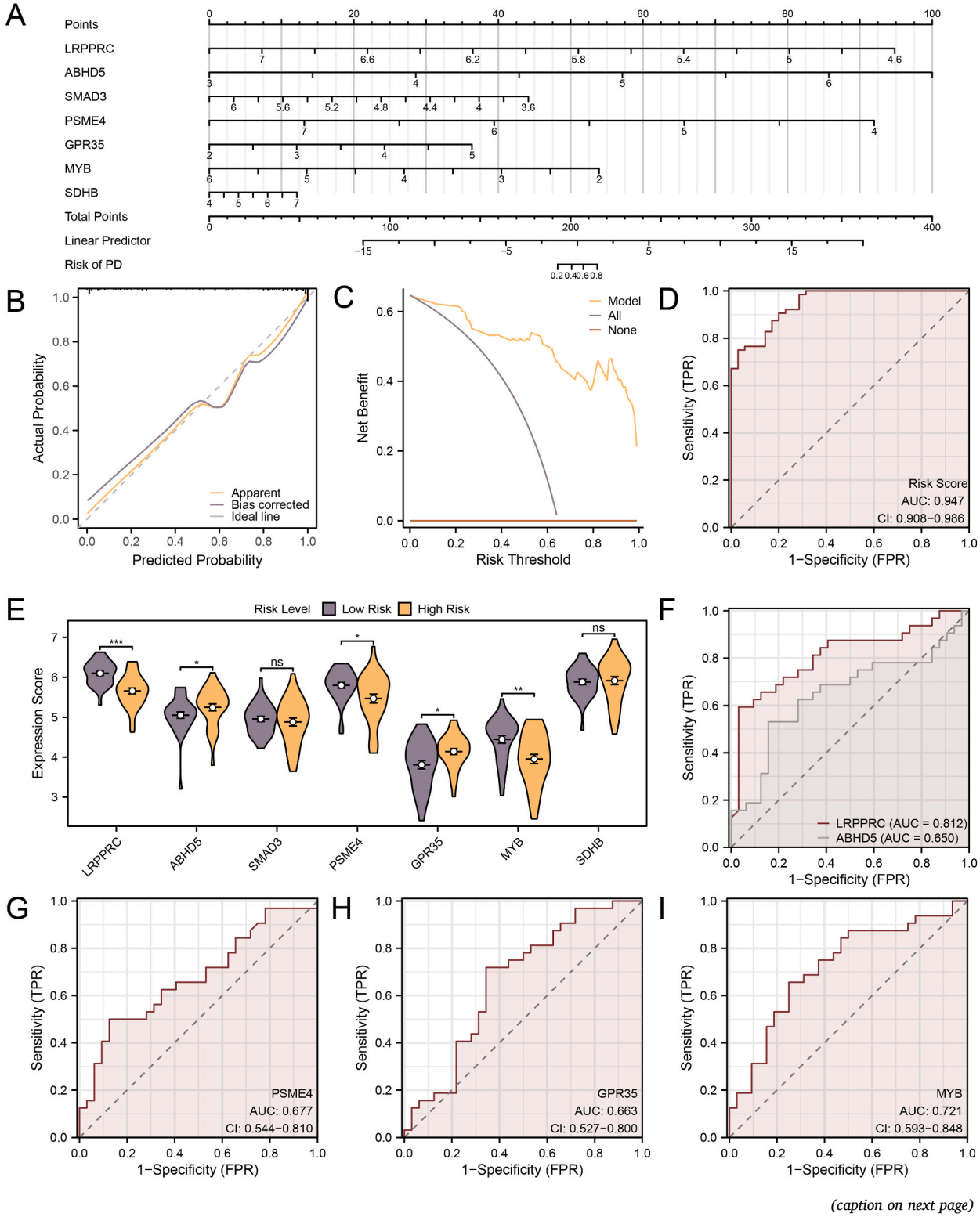
Fig. 6. Model of PD diagnostics.
A. As part of the diagnostic assessment of PD, 26 GRDEGs were plotted as forest plots. B-C. Using the SVM algorithm, genes with the lowest error rate (B) and highest accuracy (C) can be viewed. D-E. A diagnostic model plot (D) and a variable trajectory plot (E) of the LASSO regression model. PD, Parkinsons disease; GRDEGs, glycolysis related differentially expressed genes; SVM, support vector machine; LASSO, least absolute shrinkage and selection operator.

3.1.6. Validation of Parkinson's diagnostic model

A nomogram illustrating the relationships among the key genes in the merged GEO datasets (Fig. 7A) was created to validate the PD diagnostic model. These findings indicated that the ABHD5 expression level was much more useful for the PD diagnostic model than the other variables. The expression level of SDHB in the PD diagnostic model was markedly lower than that of the remaining variables.

For assessment, a calibration curve was constructed through a calibration analysis to assess the predictive performance of the PD diagnostic model based on the goodness of fit (Fig. 7B). The calibration curve of the MD diagnostic model displayed a gray line that diverged slightly from the diagonal line of the idealized model yet remained closely aligned with it, indicating consistency. DCA was used to assess the clinical usefulness of the MD diagnostic models utilizing key genes from the combined GEO datasets (Fig. 7C). The findings indicated that the trend line of the model was more consistent within a specific range than both positive and negative trends. In addition, the model yielded higher net benefits and demonstrated superior performance. The ROC curve was generated using the R package pROC to visualize the LASSO risk score in the combined GEO datasets. The following formula was used to calculate the LASSO risk score:

$$\text{risk}Score = \text{LRPPRC}^{*}\,(-1.549) + \text{ABHD5}^{*}\,(1.330) + \text{SMAD3}^{*}\,(-0.572) + \text{PSME4}^{*}\,(-0.885) + \text{GPR35}^{*}\,(0.691) + \text{MYB}^{*}\,(-0.402) + \text{SDHB}^{*}\,(0.177)$$



(caption on next page)

Fig. 7. Diagnostic analysis of PD
A. Nomograms of key genes in combined GEO datasets in PD diagnostic models. B-C. Calibration curve plot (B) and DCA plot (C) of key genes in the integrated GEO datasets for the PD diagnostic model. D. ROC curve of the LASSO risk score in the combined GEO datasets. E. Group comparison of key genes in high- and low-risk groups of PD samples. F-I. LRPPRC, ABHD5 (F), and PSME4 (G) are key genes with significant differences in expression values between the high- and low-risk groups of PD samples. ROC curves for GPR35 (H) and MYB (I) expression in PD samples. The ordinate of the calibration curve is the net benefit, and the abscissa represents the threshold probability. PD, Parkinson's disease; DCA, decision curve analysis; ROC, receiver operating characteristic; LASSO, least absolute shrinkage and selection operator. ns represents p value ≥ 0.05, no statistical significance; * represents p value < 0.05, indicating statistical significance; ** represents p value < 0.01, high statistical significance; *** represents a p value < 0.001, which indicates the highest statistical significance. AUC >0.5 indicates if the expression of the molecule promotes occurrence, and the closer it is to 1, the more effective the diagnostic effect. An AUC of 0.5–0.7 had low accuracy, 0.7 to 0.9 had moderate accuracy, and anything over 0.9, high accuracy.

Next, PD specimens were categorized according to the median LASSO risk score from the MD diagnostic model to investigate the varying expression of Key Genes in PD samples. The group comparison figure (Fig. 7E) displays the analysis results of the expression levels of seven key genes in the high-risk group compared to the low-risk group in PD samples. The findings from the comparison (Fig. 7E) indicated a highly significant difference ($p < 0.001$) in the expression of the LRPPRC between the high- and low-risk groups of PD samples. The expression levels of MYB were highly significant ($p < 0.01$), and those of key genes ABHD5, PSME4, and GPR35 were statistically significant ($p < 0.05$). Finally, the R package pROC was used to draw ROC curves of the key genes with significant differences in expression values between the high- and low-risk groups of PD samples. The ROC curve (Fig. .7F–I) showed that the expression of LRPPRC and MYB was accurate in the diagnosis of PD in the high- and low-risk groups ($0.7 < AUC < 0.9$). The expression of PSME4 and GPR35 had low accuracy (0.5 the AUC <0.7) for diagnosing PD in the high- and low-risk groups.

3.1.7. Comparing high and low expression groups through the use of differential analysis and GSEA

According to the median risk score from the LASSO in the PD diagnostic model, the PD samples were categorized into high- and low-risk categories. In the analysis of DEGs for PD using the "limma" [22] R package, a total of 461 samples met the criteria of $|\log FC| > 0.30$ and $p < 0.05$. Among these samples, 243 genes were found to be downregulated ($\log FC < -0.30$ and $p < 0.05$) while 218 genes were upregulated ($\log FC > 0.30$ and $p < 0.05$). A volcano plot was generated based on the results of the differential analysis of this dataset (Fig. 8A). The R package "pheatmap" was utilized to create a heatmap illustrating the results of the differential analysis of the DEGs as shown in Fig. 8B.

GSEA was used to examine how the expression levels of all genes in an MD sample affected the likelihood of developing MD by analyzing their involvement in BP, CC, and MF (Fig. 8C). Table 4 presents the outcomes. The findings indicated a significant enrichment of all genes in MD samples in signaling by NOTCH4 (Fig. 8D), as well as p130cas linkage to MAPK signaling for integrins (Fig. 8E). TNF receptor superfamily (TNFRSF) members are involved in the noncanonical NF-κB pathway (Fig. 8F) while TP53 controls the transcription of DNA repair genes (Fig. 8G) alongside other biological processes and signaling pathways.

3.1.8. Construction of regulatory network

Following the acquisition of miRNAs associated with the primary genes from TarBase, an mRNA-miRNA regulatory network was created and visualized using Cytoscape (Fig. 9A), and 7 key genes and 32 miRNAs were identified. Refer to Table S2 for detailed information.

Using the ChIPBase and HF target databases, we identified a combination of TFs and key genes. Using Cytoscape, we constructed and visualized an mRNA-TF regulatory network (Fig. 9B). 7 key genes and 40 TFs were included in this group, and with more details found in Table S3.

3.1.9. Immune infiltration analysis of GEO integrated dataset

the CIBERSORT algorithm was applied to the combined GEO datasets to measure the extent of immune infiltration and connect it with 22 different immune cells between the PD and control groups. As indicated by the immune infiltration analysis results, CIBERSORT immune cell proportions were calculated for the integrated GEO dataset and are shown in a bar chart (Fig. 10A). Based on the CIBERSORT immune infiltration analysis (Fig. 10B), the correlation heatmap displayed the immune cell infiltration abundance. Plasma cells and mast cells were positively correlated ($r = 0.538$), while CD8 T cells and neutrophils were negatively correlated ($r = -0.474$). Key genes and immune cells with significant correlations ($p < 0.05$) were screened and displayed using a correlation heat map (Fig. 10C). The correlation heat map showed that the key gene MYB had the strongest positive correlation with CD4 naive T cells ($r = 0.314$). The key gene SDHB showed the strongest negative correlation with naïve CD4 T cells ($r = -0.294$).

The ssGSEA algorithm was used to calculate the levels of immune infiltration by 28 immune cells in the PD and control groups. ssGSEA immune infiltration analysis revealed the correlation of immune cell infiltration abundance through correlation heat maps (Fig. 10D). Neutrophils and macrophage had the most pronounced positive correlation (r value $= 0.669$), while activated dendritic cells and activated T cells had the strongest inverse correlation (r value $= -0.443$).

Finally, key genes with a significant correlation ($p < 0.05$) and immune cells were screened and displayed using a correlation heat map (Fig. 10E). A correlation heat map showed that ABHD5 correlated most positively with neutrophils (r value $= 0.525$), while SMAD3 had the greatest negative correlation with mast cells (r value $= -0.250$).

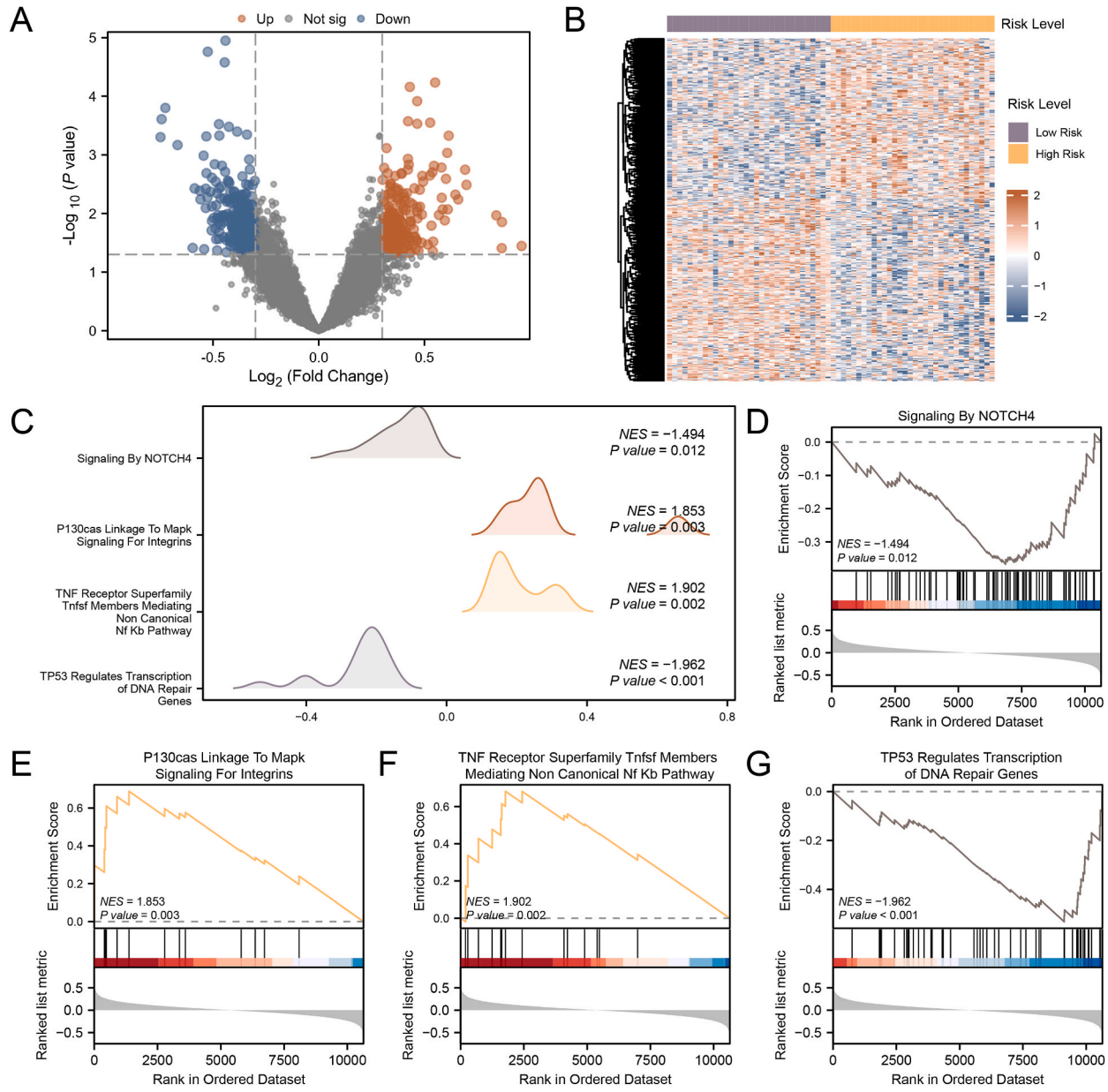


Fig. 8. Differential gene expression analysis and GSEA by risk group.
A. PD samples with a volcano plot of differentially expressed genes. B. An illustration of the levels of gene expression in samples from PD patients at different risk. C. Display of 4 biological functions of GSEA mountain map of PD samples. GSEA showed that those at high risk for PD were significantly enriched in signaling by NOTCH4 (D), and p130cas linkage to MAPK signaling for integrins (E). TNF receptor superfamily TNFSF members underlying the noncanonical NF-κB pathway (F) and TP53 regulates transcription of DNA repair genes (G). PD, Parkinson's disease; GSEA, gene set enrichment analysis. The screening criteria of the GSEA were p value < 0.05 and FDR value (q value) < 0.25.

3.1.10. Immune infiltration analysis of Parkinson's samples

The CIBERSORT algorithm was used with combined GEO datasets to determine the abundance of immune infiltration and the correlation among 22 immune cells in the high- and low-risk groups of PD samples. A bar graph illustrating the percentage of CIBERSORT immune cells in the combined GEO datasets was created based on the immune infiltration analysis results (Fig. 11A). The correlation heat map displayed the results of the CIBERSORT immune infiltration analysis (Fig. 11B). The results showed that the strongest positive correlation was between M1 macrophages and resting dendritic cells ($r = 0.388$). The strongest negative correlation was found between CD8 T cells and neutrophils (r value $= -0.538$). Subsequently, key genes with a significant correlation ($p < 0.05$) and immune cells were screened and displayed using a correlation heat map (Fig. 11C). Based on the correlation heat map, the key gene ABHD5 correlated the strongest with neutrophils (r value $= -0.536$). ABHD5 level was positively correlated with CD8 T cells ($r =$

Table 4
Results of GSEA for risk group.

ID	Set Size	Enrichment Score	NES	pvalue	p.adjust	qvalue
WP_EICOSANOID_SYNTHESIS	19	0.727038	2.205316	1.78 e−05	0.001828	0.001624
BIOCARTA_EICOSANOID_PATHWAY	22	0.688873	2.179429	0.000113	0.007197	0.006393
KEGG_ARACHIDONIC_ACID_METABOLISM	46	0.569975	2.131868	2.6 e−05	0.002562	0.002276
WP_PROSTAGLANDIN_AND_LEUKOTRIENE_METABOLISM_IN_SENESCENCE	29	0.620087	2.094987	7.84 e−05	0.005614	0.004987
REACTOME_RESPONSE_TO_ELEVATED_PLATELET_CYTOSOLIC_CA2	111	0.461161	2.042921	1.67 e−06	0.000219	0.000194
REACTOME_RUNX1_REGULATES_GENES_INVOLVED_IN_MEGAKARYOCYTE_DIFFERENTIATION_AND_PLATELET_FUNCTION	19	0.671524	2.036928	0.000299	0.015712	0.013956
WP_PLATELETMEDIATED_INTERACTIONS_WITH_VASCULAR_AND_CIRCULATING_CELLS	16	0.701543	2.026058	0.000697	0.026575	0.023605
PID_S1P_S1P4_PATHWAY	13	0.737962	1.984987	0.000464	0.019783	0.017572
REACTOME_PLATELET_ACTIVATION_SIGNALING_AND_AGGREGATION	227	0.401373	1.947741	5.69 e−08	9.16 e−06	8.13 e−06
WP_SELENIUM_MICRONUTRIENT_NETWORK	61	0.481255	1.916518	0.000309	0.015855	0.014083
REACTOME_TNF_RECEPTOR_SUPERFAMILY_TNFSF_MEMBERS_MEDIATING_NON_CANONICAL_NF_KB_PATHWAY	14	0.682362	1.902147	0.002143	0.056267	0.04998
REACTOME_GRB2_SOS_PROVIDES_LINKAGE_TO_MAPK_SIGNALING_FOR_INTEGRINS	14	0.682205	1.901711	0.002143	0.056267	0.04998
WP_VITAMIN_D_RECEPTOR_PATHWAY	142	0.416697	1.899787	6.64 e−06	0.000785	0.000697
WP_CHOLESTASIS	17	0.643998	1.877119	0.002672	0.064425	0.057225
PID_EPHRINB_REV_PATHWAY	29	0.554653	1.873917	0.001279	0.040312	0.035807
REACTOME_SMOOTH_MUSCLE_CONTRACTION	36	0.524145	1.862753	0.001392	0.042158	0.037447
PID_EPO_PATHWAY	32	0.541318	1.859103	0.000834	0.030792	0.027351
REACTOME_P130CAS_LINKAGE_TO_MAPK_SIGNALING_FOR_INTEGRINS	13	0.688808	1.852773	0.003052	0.070027	0.062202
REACTOME_SIGNALING_BY_NOTCH4	75	0.36647	1.49361	0.011631	0.148548	0.131949
REACTOME_TP53_REGULATES_TRANSCRIPTION_OF_DNA_REPAIR_GENES	47	0.5314	1.96208	8.38 e−05	0.005725	0.005085

GSEA, Gene Set Enrichment Analysis.

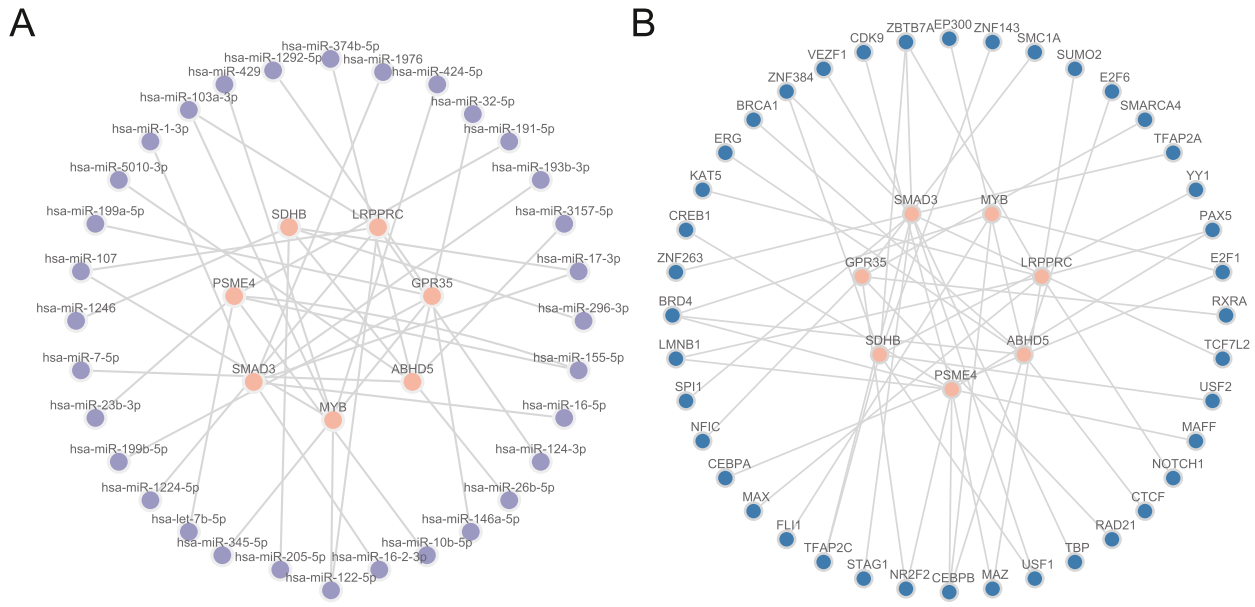


Fig. 9. Regulatory network of key genes.
A. The mRNA-miRNA regulatory network of key genes. B. mRNA-TF regulatory network of key genes. TF: transcription factor. Orange, purple, and blue circles represent mRNAs, miRNAs, and transcription factors, respectively.

−0.488).

Simultaneously, the ssGSEA algorithm was used to calculate the quantity and relationship of immune infiltration from 28 immune cells in both the high- and low-risk groups of PD samples. First, the correlation heat maps (Fig. 11D) displayed the results of immune cell infiltration abundance in the ssGSEA immune infiltration analysis. The most robust positive correlation was observed between neutrophils and macrophages ($r = 0.660$). Immature dendritic cell and activated CD4 T cell (r value = −0.522) showed the strongest negative correlation.

Finally, we screened key genes with a significant correlation with immune cells ($p < 0.05$) and presented them using a correlation heat map (Fig. 11E). The correlation heatmap showed the strongest positive correlation between ABHD5 and neutrophils ($r = 0.507$) among the key genes. The key gene SMAD3 showed the strongest negative correlation with gamma delta T cells (r value = −0.488).

4. Discussion

PD is a neurodegenerative condition that severely affects the quality of life of patients as the disease progresses. Therefore, exploring the molecular mechanisms and risk prediction models of PD is a valuable endeavor. The relationship between glycolysis and PD is a topic of discussion in the literature. Research has indicated that there is a notable reduction in glycolysis levels in the cerebral cortex of individuals with PD compared to the general population [41]. Increasing glycolysis to boost ATP production alleviates certain symptoms in patients with PD [42]. By understanding how glycolytic pathways can be leveraged to compensate for mitochondrial deficits, we can propose novel interventions to promote neuronal survival and function. For example, Tretter and Adam-Vizi illustrated that enhancing glycolytic metabolism could create a balanced redox environment, lowering oxidative stress and potentially slowing neurodegeneration [43]. Such insights suggest that targeted metabolic therapies, aimed at boosting glycolysis or restoring mitochondrial function, could provide a dual benefit in combating the effects of PD-associated mitochondrial impairment.

In this study, glycolysis was used as the starting point to identify a gene regulatory network associated with PD risk. Through a series of bioinformatics analyses, seven key genes associated with PD were identified (LRPPRC, ABHD5, SMAD3, PSME4, GPR35, MYB, and SDHB). Some target miRNAs and transcription factors that might be related to PD were also predicted to find new directions for further investigation of the molecular mechanisms of PD.

To further determine the value of GRDEGs in the diagnosis of PD, we constructed and validated a PD diagnostic model. The model was constructed using the SVM algorithm and achieved the highest accuracy with a gene count of seven GRDEGs: LRPPRC, ABHD5, SMAD3, PSME4, GPR35, MYB, and SDHB. A LASSO regression model was further constructed, and the seven GRDEGs were LRPPRC, ABHD5, SMAD3, PSME4, GPR35, MYB, and SDHB. After additional verification and analysis of the diagnostic model, it was discovered that the expression level of the key gene ABHD5 had a more pronounced impact on the PD diagnostic model than other factors. Key gene LRPPRC in PD samples showed significant variation between the high- and low-risk groups ($p < 0.001$), with highly significant differences in expression ($p < 0.01$) and significant differences in expression levels ($p < 0.05$). The difference in the expression of the key gene MYB was highly statistically significant ($p < 0.01$), and the expression levels of the key genes ABHD5, PSME4, and GPR35 were significantly different ($p < 0.05$). Among these, LRPPRC maintains mitochondrial potential and activity, and high levels of

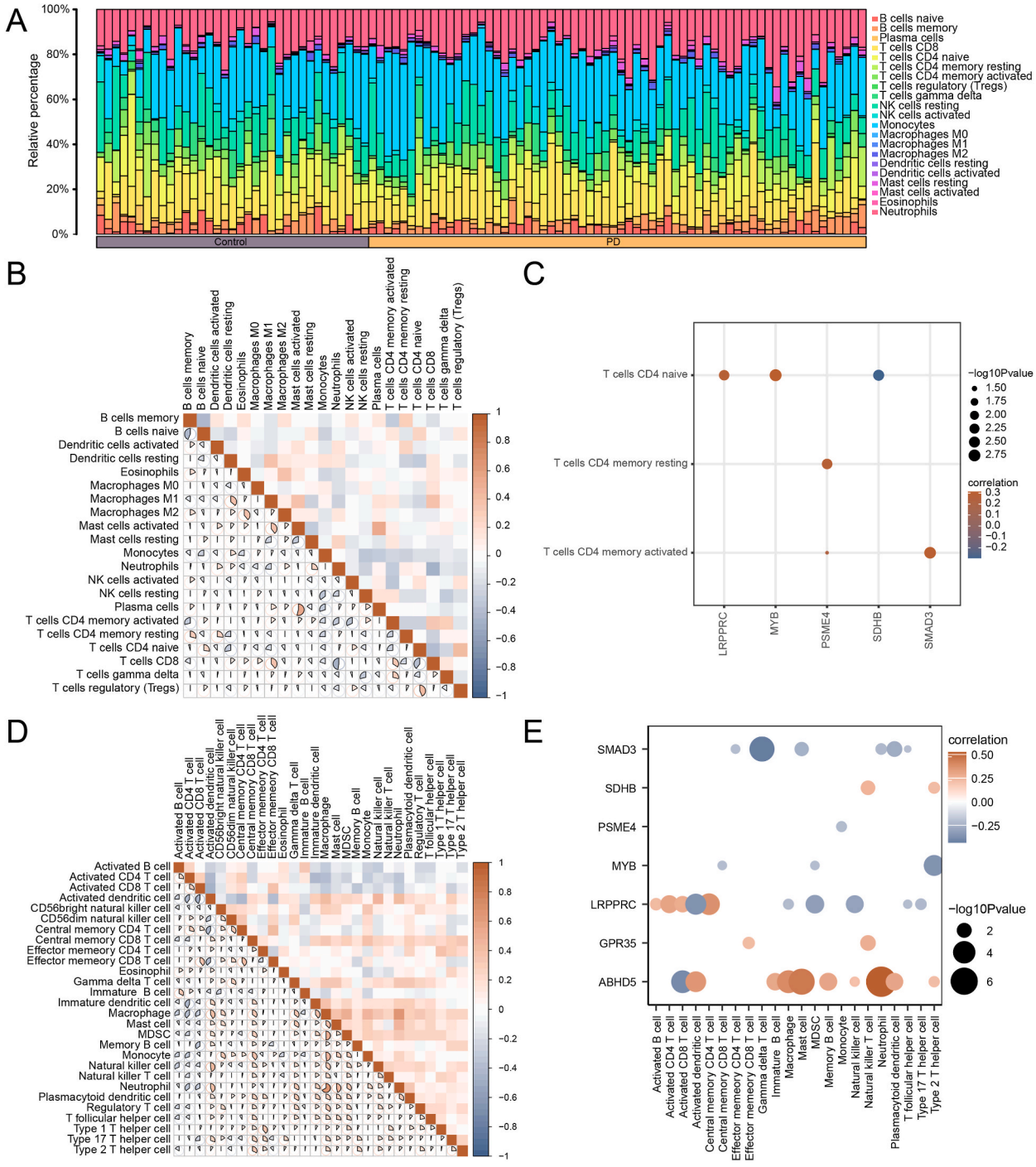
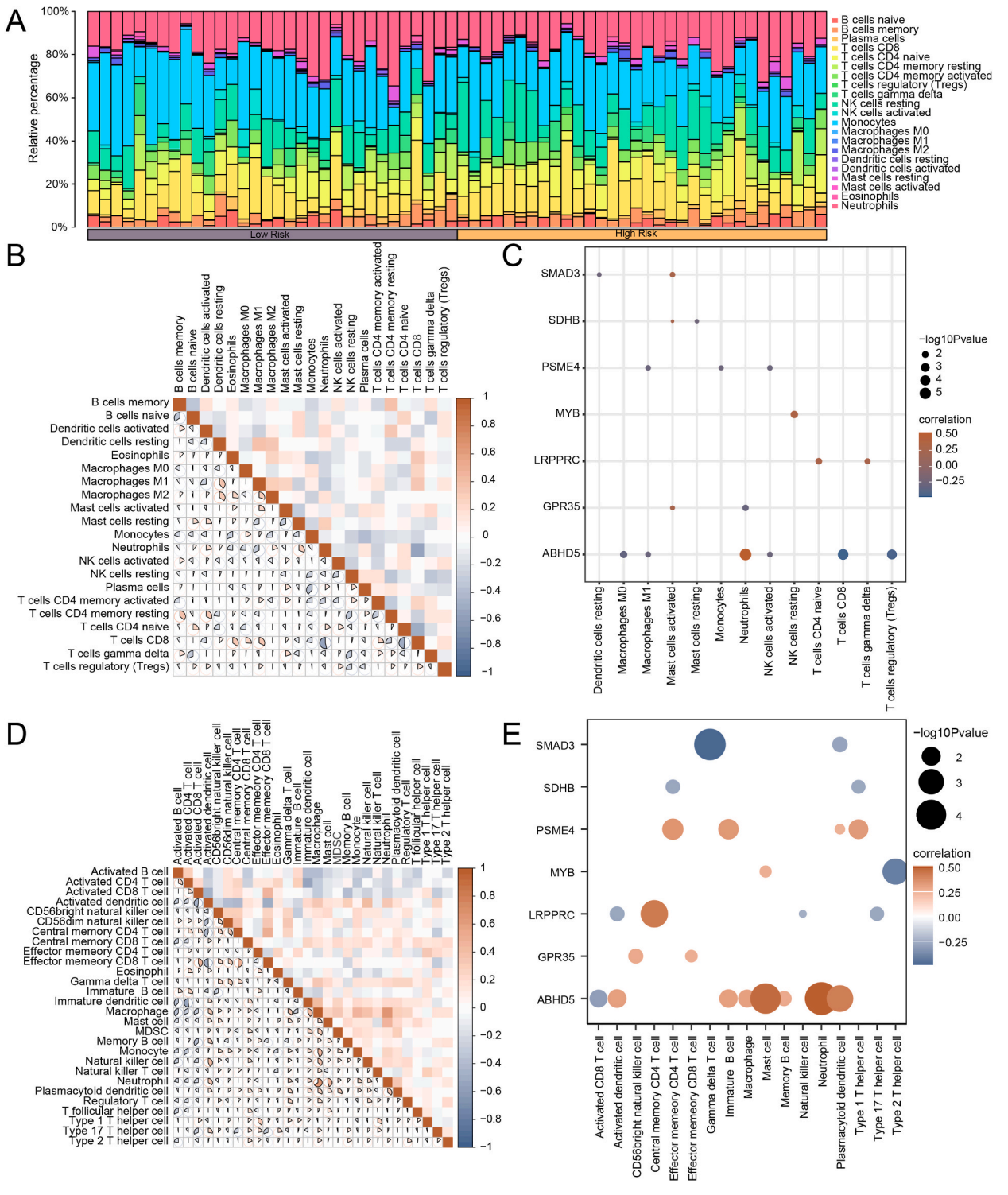


Fig. 10. Infiltration of immune cells in combined datasets.

A.CIBERSORT immune cell infiltration abundance proportion bar graph of combined GEO datasets. B. Correlation heatmap of CIBERSORT immune cell infiltration abundance in integrated GEO datasets (combined datasets). C. Heat map of the correlation between key genes and CIBERSORT immune cell infiltration abundance in the combined GEO datasets. D. Correlation heatmap of ssGSEA immune cell infiltration in integrated GEO datasets (combined datasets). E. The infiltration abundance of ssGSEA immune cells in a heat map of key genes and the combined datasets. PD, Parkinson's disease; ssGSEA, single-sample gene set enrichment analysis. Absolute values of the correlation coefficient (r value) below 0.3 indicated weak or no correlation, between 0.3 and 0.5 was weak correlation, and between 0.5 and 0.8 was moderate correlation. In the proportion bar chart, orange represents the PD group, and purple represents the control group. The correlation heat map shows the positive and negative correlations (red and blue, respectively).



(caption on next page)

Fig. 11. Analyses of immune infiltration in risk groups.
A. CIBERSORT immune cells infiltration abundance proportion bar chart in PD samples. B. Correlation heatmap of CIBERSORT immune cell infiltration abundance in PD samples. C. Key genes and CIBERSORT immune cell abundance in PD samples on a heat map. D. Correlation heatmap of immune cell infiltration abundance of ssGSEA in PD samples. E. In PD samples, a heat map correlates key genes with ssGSEA immune cell infiltration abundance. PD, Parkinson's disease; ssGSEA, single-sample gene-set enrichment analysis. An absolute value of the correlation coefficient (r value) below 0.3 indicated weak or no correlation, between 0.3 and 0.5 was weak correlation, and between 0.5 and 0.8 was moderate correlation. In the proportion bar chart, orange is the high-risk group, and purple is the control group. In the correlation heat map, red is the positive correlation and blue is the negative correlation.

LRPPRC suggest higher mitochondrial activity, which is related to the glycolytic pathway [44]. ABHD5 is involved in fat metabolism and might also be involved in glycolytic pathways [45].

The alpha-beta hydrolase domain-containing protein 5 (ABHD5) functions as an activating coenzyme for adipose triglyceride lipase (ATGL), the primary lipase in several tissues [46]. A previous study has shown that ABHD5 prevents aerobic glycolysis and reinforces the ability of PTEN and p53 to suppress tumors [47]. Therefore, ABHD5 prevents glycolysis. This study indicates that ABHD5 in the PD diagnostic model has an impact on the glycolytic pathway, potentially influencing the development of PD. However, further research is required to determine whether this effect involves the inhibition or promotion of glycolysis.

The RNA-binding protein LRPPRC is responsible for regulating mRNA stability and polyadenylation, mainly within the mitochondria, but also in the cytoplasm and nucleus [48]. In addition to its role in mitochondrial translation and RNA decomposition, LRPPRC deficiency affects the mitochondrial electron transport chain function, mitochondrial permeability, and transmembrane ROS diffusion [49]. In our study, the expression of LRPPRC was negatively correlated with the risk of PD, suggesting that LRPPRC may be involved in the pathogenesis of PD by destroying mitochondrial stability; however, no relevant reports have been published.

Recent studies have focused on the immunological microenvironment of PD. It has been shown that inflammation plays an important role in the progression of neurodegenerative diseases, and that immune cells (such as microglia and peripheral immune cells) may contribute to neuronal damage and disease deterioration [50]. There is evidence of immunological abnormalities associated with PD as well as a range of diseases with heterogeneous etiologies. Patients with PD show a wide range of peripheral immune abnormalities, including a general reduction in lymphocyte numbers, decreases in CD19 B, and decreases in CD3 and CD4 subsets [51]. In this study, we analyzed the expression levels of PD-associated immune cells and found that the most strongly associated immune cells were CD8 T cells and neutrophils, which negatively correlated with ABHD5 expression. This study provides a new approach for exploring the pathogenesis of PD.

In PD, GRDEGs were mainly enriched in regulation of protein export from the nucleus, positive regulation of proteolysis, activation of cysteine-type endopeptidases involved in apoptosis, and regulation of lipid localization, BP such as nuclear export and protein processing; the lumens of secretory granules, cytoplasmic vesicles, and vesicles, nuclear inner membrane, ficolin-1-rich granule and nuclear membrane components; peptidase activator activity, phosphatase binding, glucose binding, cysteine-type endopeptidase activity, cysteine-type endopeptidase activator activity is involved in apoptotic process, monosaccharide binding and other molecular functions. Furthermore, it was enriched in the glucagon signaling and apoptosis pathways. However, we need to note that the glycolysis mechanism and function between blood and nerve cells are different. First, there are some differences in the glycolytic pathways of blood cells and nerve cells. Blood cells rely primarily on aerobic glycolysis for energy production, while nerve cells rely primarily on anaerobic glycolysis. This is because nerve cells need a quick supply of energy to perform their functions, and anaerobic glycolysis is able to provide energy more quickly. Secondly, there are also differences in the metabolites produced by blood cells and nerve cells during glycolysis. Blood cells mainly produce lactic acid, while nerve cells mainly produce pyruvate. Changes in the concentration of these metabolites inside and outside the cell affect the energy supply and metabolic balance of the cell. In addition, abnormalities in the glycolysis of blood cells and nerve cells can lead to different diseases. In blood cells, dysglycolysis may lead to hyperglycemia and hyperinsulinemia; In nerve cells, glycolysis disorders may lead to neurodegenerative diseases such as Alzheimer's disease. Therefore, in subsequent studies, we will further explore the role of glycolysis on nerve cells and PD.

Previous studies have indicated that $CD4^+$ cells play an important role in the progression of PD [41] and that SDHB is highly associated with mitochondrial dysfunction [52]. Hence, exploring the relationship between SDHB and neuroinflammation in PD may provide new insights into PD development.

Additionally, metabolic alterations that arise from these processes can have a profound impact on neuronal health. Dysregulated glycolysis not only influences the energy supply for neuronal function but also increases oxidative stress, a well-recognized contributor to neurodegeneration. The consequence of impaired energy metabolism leads to neuronal apoptosis and loss of synaptic integrity, thus worsening the symptoms associated with PD. Understanding how these interrelated processes contribute to neuronal and systemic alterations is essential for developing targeted therapeutic strategies aimed at mitigating both metabolic disturbances and neuroinflammation.

This study has some limitations. First, it lacked important clinical information, including details of disease activity and medication use. For example, sex and age are important factors affecting many biological processes, including changes in the expression of glycolytic pathways. However, due to the lack of gender and age information in the GSE6613 and GSE54536 datasets, we were unable to directly investigate the effect of these factors on the expression changes of glycolytic pathways and their association with PD, which may limit our understanding of this issue. Second, is the type of sample used in this study. Although we analyzed gene expression changes in blood samples from patients with Parkinson's disease, we did not consider analysis of nerve tissue samples. Gene expression signatures in blood samples may not be fully representative of neurological alterations in the disease process, which may affect the comprehensiveness and accuracy of our findings. In addition, our study was based on a bioinformatic approach to initially explore the

diagnostic model of Parkinson's disease and its potential diagnostic capabilities, which meant that sample source was not considered as a primary consideration in the study. Therefore, future studies need to incorporate neural tissue samples to further validate our findings and improve the validity and applicability of the diagnostic model. Because our study did not carefully distinguish the contribution and activity status of each type of cell, our analysis may not fully reflect the true biology of the disease. Future studies should consider the composition of different cell types in the blood and their effects on gene expression to improve the reliability of the results and the accuracy of clinical applications. Finally, future research should not only focus on PD but also consider the role of glycolysis in other age-related neurodegenerative disorders such as Alzheimer's disease (AD). Recent studies have begun to reveal that altered glucose metabolism is a common feature of various neurodegenerative diseases, suggesting that the pathways involved in glycolysis may be broadly implicated in neuronal dysfunction across multiple contexts [53]. For instance, understanding how glucose hypometabolism is associated with cognitive decline in Alzheimer's could unveil shared mechanisms with PD, potentially revealing overlapping metabolic targets for therapeutic interventions.

5. Conclusion

Our study provides valuable information on the glycolytic landscape associated with PD and identifies reliable diagnostic biomarkers of this disease. LRPPRC, ADHB, CD8 T cells, and CD4 T cells may participate in the development of PD. Therefore, this study provides a new perspective on the diagnosis and molecular mechanisms of PD.

CRediT authorship contribution statement

Jing Shen: Writing – original draft, Methodology, Data curation, Conceptualization. **Ensheng Yao:** Data curation, Conceptualization. **Weidong Tian:** Validation, Data curation. **Jia He:** Writing – review & editing, Methodology, Conceptualization. **Yukai Gu:** Data curation. **Dong Zhao:** Writing – review & editing, Conceptualization.

Funding

Tianshan Talent Platform construction project - Clinical Center for Neurological diseases (CZ001210)
Postdoctoral Foundation of Xinjiang Production and Construction Corps (No. 338594).
Guiding Science and Technology Program Projects of the XinJiang Corps (No.2024ZD049).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors thank all the people and who participated in this study, and their caregivers who facilitated their participation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2025.e41831.

References

[1] R. Balestrino, A.H.V. Schapira, Parkinson disease, Eur. J. Neurol. 27 (1) (2020) 27–42.
[2] A. Ascherio, M.A. Schwarzschild, The epidemiology of Parkinson's disease: risk factors and prevention, Lancet Neurol. 15 (12) (2016) 1257–1272.
[3] S.Y. Lunt, M.G. Vander Heiden, Aerobic glycolysis: meeting the metabolic requirements of cell proliferation, Annu. Rev. Cell Dev. Biol. 27 (2011) 441–464.
[4] A. Samii, J.G. Nutt, B.R. Ransom, Parkinson's disease, Lancet 363 (9423) (2004) 1783–1793.
[5] R. Cacabelos, Parkinson's disease: from pathogenesis to pharmacogenomics, Int. J. Mol. Sci. 18 (3) (2017).
[6] A.H.H. Siddique, P.P. Kale, Importance of glucose and its metabolism in neurodegenerative disorder, as well as the combination of multiple therapeutic strategies targeting α-synuclein and neuroprotection in The Treatment of Parkinson's Disease, (35-3787 (Print)).
[7] T. Briston, A.R. Hicks, Mitochondrial dysfunction and neurodegenerative proteinopathies: mechanisms and prospects for therapeutic intervention, Biochem. Soc. Trans. 46 (4) (2018) 829–842.
[8] A.M. Smith, C. Depp, B.J. Ryan, G.I. Johnston, J. Alegre-Abarrategui, S. Evetts, M. Rolinski, F. Baig, C. Ruffmann, A.K. Simon, M.T.M. Hu, R. Wade-Martins, Mitochondrial dysfunction and increased glycolysis in prodromal and early Parkinson's blood cells, Mov. Disord. 33 (10) (2018) 1580–1590.
[9] S.R. Subramaniam, M.F. Chesselet, Mitochondrial dysfunction and oxidative stress in Parkinson's disease, Prog. Neurobiol. 106–107 (2013) 17–32.
[10] U. Naeem, A.R. Arshad, A. Jawed, F. Eqbal, L. Imran, Z. Khan, F. Ijaz, Glycolysis: the next big breakthrough in Parkinson's disease, Neurotox. Res. 40 (6) (2022) 1707–1717.
[11] B.L. Tang, Glucose, glycolysis, and neurodegenerative diseases, J. Cell. Physiol. 235 (11) (2020) 7653–7662.
[12] P.A. Dionisio, J.D. Amaral, C.M.P. Rodrigues, Oxidative stress and regulated cell death in Parkinson's disease, Ageing Res. Rev. 67 (2021) 101263.
[13] J. Li, L. Chen, Q. Qin, D. Wang, J. Zhao, H. Gao, X. Yuan, J. Zhang, Y. Zou, Z. Mao, Y. Xiong, Z. Min, M. Yan, C.Y. Wang, Z. Xue, Upregulated hexokinase 2 expression induces the apoptosis of dopaminergic neurons by promoting lactate production in Parkinson's disease, Neurobiol. Dis. 163 (2022) 105605.

[14] Y. Miki, S. Shimoyama, T. Kon, T. Ueno, R. Hayakari, K. Tanji, T. Matsumiya, E. Tsushima, F. Mori, K. Wakabayashi, M. Tomiyama, Alteration of autophagy-related proteins in peripheral blood mononuclear cells of patients with Parkinson's disease, Neurobiol. Aging 63 (2018) 33–43.

[15] C.R. Scherzer, A.C. Eklund, L.J. Morse, Z. Liao, J.J. Locascio, D. Fefer, M.A. Schwarzschild, M.G. Schlossmacher, M.A. Hauser, J.M. Vance, L.R. Sudarsky, D. G. Standaert, J.H. Growdon, R.V. Jensen, S.R. Gullans, Molecular markers of early Parkinson's disease based on gene expression in blood, Proc. Natl. Acad. Sci. U.S.A. 104 (3) (2007) 955–960.

[16] C.R. Scherzer, J.A. Grass, Z. Liao, I. Pepivani, B. Zheng, A.C. Eklund, P.A. Ney, J. Ng, M. McGoldrick, B. Mollenhauer, E.H. Bresnick, M.G. Schlossmacher, GATA transcription factors directly regulate the Parkinson's disease-linked gene alpha-synuclein, Proc. Natl. Acad. Sci. U.S.A. 105 (31) (2008) 10907–10912.

[17] A. Alieva, M.I. Shadrina, E.V. Filatova, A.V. Karabanov, S.N. Illarioshkin, S.A. Limborska, P.A. Slominsky, 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.

[18] M.P. Davis S, GEOquery: a bridge between the gene expression Omnibus (GEO) and BioConductor, Bioinformatics 23 (14) (2007 Jul 15) 1846–1847.

[19] G. Stelzer, N. Rosen, I. Plaschkes, S. Zimmerman, M. Twik, S. Fishilevich, T.I. Stein, R. Nudel, I. Lieder, Y. Mazor, S. Kaplan, D. Dahary, D. Warshawsky, Y. Guan-Golan, A. Kohn, N. Rappaport, M. Safran, D. Lancet, The GeneCards suite: from gene data mining to disease genome sequence analyses, Curr Protoc Bioinformatics 54 (2016) 1.30.1–1.30.33.

[20] A. Liberzon, A. Subramanian, R. Pinchback, H. Thorvaldsdóttir, P. Tamayo, J.P. Mesirov, Molecular signatures database (MSigDB) 3.0, Bioinformatics 27 (12) (2011) 1739–1740.

[21] J.T. Leek, W.E. Johnson, H.S. Parker, A.E. Jaffe, J.D. Storey, The sva package for removing batch effects and other unwanted variation in high-throughput experiments, Bioinformatics 28 (6) (2012) 882–883.

[22] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47.

[23] H. Mi, A. Muruganujan, D. Ebert, X. Huang, P.D. Thomas, PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools, Nucleic Acids Res. 47 (D1) (2019) D419–d426.

[24] M. Ashburner, J.A. Ball Ca Fau - Blake, D. Blake Ja Fau - Botstein, H.Botstein D. Fau - Butler, J.M. Butler, H. Fau - Cherry, A.P. Cherry Jm Fau - Davis, K. Davis Ap Fau - Dolinski, S.S.Dolinski K. Fau - Dwight, J.T. Dwight Ss Fau - Eppig, M.A. Eppig Jt Fau - Harris, D.P. Harris Ma Fau - Hill, L. Hill Dp Fau - Issel-Tarver, A. Issel-Tarver L Fau - Kasarskis, S.Kasarskis A. Fau - Lewis, J.C. Lewis, S. Fau - Matese, J.E. Matese Jc Fau - Richardson, M. Richardson Je Fau - Ringwald, G. M. Ringwald M Fau - Rubin, G. Rubin Gm Fau - Sherlock, G. Sherlock, Gene ontology: tool for the unification of biology, The Gene Ontology Consortium 25 1 (2000) 25–29.

[25] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (1) (2000) 27–30.

[26] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (5) (2012) 284–287.

[27] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U. S. A. 102 (43) (2005) 15545–15550.

[28] G. Yu, Y. Wang Lg Fau - Han, Q.-Y. Han, Y. Fau - He, Q.Y. He, clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters, vol. 16, OMICS, 2012, pp. 1557–8100 (Electronic)).

[29] H. Sanz, C. Valim, E. Vegas, J.M. Oller, F. Reverter, SVM-RFE: selection and visualization of the most relevant features through non-linear kernels, BMC Bioinf. 19 (1) (2018) 432.

[30] B. Van Calster, L. Wynants, J.F.M. Verbeek, J.Y. Verbakel, E. Christodoulou, A.J. Vickers, M.J. Roobol, E.W. Steyerberg, Reporting and interpreting decision curve analysis: a guide for investigators, Eur. Urol. 74 (6) (2018) 796–804.

[31] F. Campo-Paysaa, M. Sémon, R.A. Cameron, K.J. Peterson, M. Schubert, microRNA complements in deuterostomes: origin and evolution of microRNAs, Evol. Dev. 13 (1) (2011) 15–27.

[32] I.S. Vlachos, M.D. Paraskevopoulou, D. Karagkouni, G. Georgakilas, T. Vergoulis, I. Kanellos, I.L. Anastasopoulos, S. Maniou, K. Karathanou, D. Kalfakakou, A. Fevgas, T. Dalamagas, A.G. Hatzigeorgiou, DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions, Nucleic Acids Res. 43 (Database issue) (2015) D153–D159.

[33] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (11) (2003) 2498–2504.

[34] K.R. Zhou, S. Liu, W.J. Sun, L.L. Zheng, H. Zhou, J.H. Yang, L.H. Qu, ChIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data, Nucleic Acids Res. 45 (D1) (2017) D43–d50.

[35] Q. Zhang, W. Liu, H.-M. Zhang, G.-Y. Xie, Y.-R. Miao, M. Xia, A.-Y. Guo, hTFtarget: a comprehensive database for regulations of human transcription factors and their targets, Dev. Reprod. Biol. 18 (2) (2020) 120–128.

[36] P. Shannon, O. Markiel A Fau - Ozier, N.S. Ozier O Fau - Baliga, J.T. Baliga Ns Fau - Wang, D. Wang Jt Fau - Ramage, N. Ramage D Fau - Amin, B. Amin N Fau - Schwikowski, T. Schwikowski B Fau - Ideker, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Print 13 11 (2003) 1088–9051, 2498-504.

[37] A.M. Newman, C.L. Liu, M.A.-O. Green, A.A.-O. Gentles, W. Feng, Y. Xu, C.D. Hoang, M. Diehn, A.A.-O. Alizadeh, Robust Enumeration of Cell Subsets from Tissue Expression Profiles, 12 5, 2015, pp. 1548–7105 (Electronic)).

[38] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, C.D. Hoang, M. Diehn, A.A. Alizadeh, Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (5) (2015) 453–457.

[39] G. Bindea, B. Mlecnik, M. Tosolini, A. Kirilovsky, M. Waldner, A.C. Obenauf, H. Angell, T. Fredriksen, L. Lafontaine, A. Berger, Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer, Immunity 39 (4) (2013) 782–795.

[40] G. Nyamundanda, P. Poudel, Y. Patil, A. Sadanandam, A Novel Statistical Method to Diagnose, Quantify and Correct Batch Effects in Genomic Studies, 7 1, 2017, pp. 2045–2322 (Electronic)).

[41] N. Kustrimovic, C. Comi, L. Magistrelli, E. Rasini, M. Legnaro, R. Bombelli, I. Aleksic, F. Blandini, B. Minafra, G. Riboldazzi, A. Sturchio, M. Mauri, G. Bono, F. Marino, M. Cosentino, Parkinson's disease patients have a complex phenotypic and functional Th1 bias: cross-sectional studies of CD4+ Th1/Th2/T17 and Treg in drug-naive and drug-treated patients, J. Neuroinflammation 15 (1) (2018) 205.

[42] S. Sakaguchi, Naturally arising CD4+ regulatory t cells for immunologic self-tolerance and negative control of immune responses, Annu. Rev. Immunol. 22 (2004) 531–562.

[43] L. Tretter, V. Sipos I Fau - Adam-Vizi, V. Adam-Vizi, Initiation of Neuronal Damage by Complex I Deficiency and Oxidative Stress in Parkinson's Disease, (364-3190 (Print)).

[44] H. Yu, Q. Chang, T. Sun, X. He, L. Wen, J. An, J. Feng, Y. Zhao, Metabolic reprogramming and polarization of microglia in Parkinson's disease: role of inflammasome and iron, Ageing Res. Rev. 90 (2023) 102032.

[45] J.M. Caviglia, D.-H. Betters Jl Fau - Dapito, C.C. Dapito Dh Fau - Lord, S. Lord Cc Fau - Sullivan, S. Sullivan S Fau - Chua, T. Chua S Fau - Yin, A. Yin T Fau - Sekowski, H. Sekowski A Fau - Mu, L. Mu H Fau - Shapiro, J.M. Shapiro L Fau - Brown, D.L. Brown Jm Fau - Brasaemle, D.L. Brasaemle, Adipose-selective Overexpression of ABHD5/CGI-58 Does Not Increase Lipolysis or Protect against Diet-Induced Obesity, (1539-7262 (Electronic)).

[46] R. Zechner, FAT FLUX: enzymes, regulators, and pathophysiology of intracellular lipolysis (2015) 1757–4684 (Electronic)).

[47] G. Chen, G. Zhou, S. Aras, Z. He, S. Lucas, I. Podgorski, W. Skar, J.G. Granneman, J. Wang, Loss of ABHD5 promotes the aggressiveness of prostate cancer cells, (2045-2322 (Electronic)).
[48] L. Volpon, B. Culjkovic-Kraljacic, H.S. Sohn, A. Blanchet-Cohen, M.J. Osborne, K.L.B. Borden, A biochemical framework for eIF4E-dependent mRNA export and nuclear recycling of the export machinery (2017) 1469–9001 (Electronic)).
[49] Y. Yao, L. Luo, G. Xiang, J. Xiong, N. Ke, C. Tan, Y. Chen, X. Liu, The Expression of M(6)A Regulators Correlated with the Immune Microenvironment Plays an Important Role in the Prognosis of Pancreatic Ductal Adenocarcinoma, (2227-684X (Print)).
[50] N. Ebadpour, M. Mahmoudi, R. Kamal Kheder, M. Abavisani, Z. Baridjavadi, N. Abdollahi, S.A. Esmaeili, From Mitochondrial Dysfunction to Neuroinflammation in Parkinson's Disease: Pathogenesis and Mitochondrial Therapeutic Approaches, (1878-1705 (Electronic)).
[51] G. Morris, M. Berk, The many roads to mitochondrial dysfunction in neuroimmune and neuropsychiatric disorders, BMC Med. 13 (2015) 68.
[52] M.I. Carlo, A.A. Hakimi, G.D. Stewart, G. Bratslavsky, J. Brugarolas, Y.B. Chen, W.M. Linehan, E.R. Maher, M.J. Merino, K. Offit, V.E. Reuter, B. Shuch, J.A. Coleman, Familial Kidney Cancer: Implications of New Syndromes and Molecular Insights, (1873-7560 (Electronic)).
[53] E. Llanos-Gonzalez, A.A. Henares-Chavarino, C.M. Pedrero-Prieto, S. Garcia-Carpintero, J. Frontinan-Rubio, F.J. Sancho-Bielsa, F.J. Alcain, J.R. Peinado, Y. Rabanal-Ruiz, M. Duran-Prado, Interplay between mitochondrial oxidative disorders and proteostasis in Alzheimer's disease, Front. Neurosci. 13 (2019) 1444.