

RESEARCH ARTICLE

Identification of cuproptosis-related key genes and pathways in Parkinson's disease via bioinformatics analysis

Jia Song, Jia Li, Xiaochen Pei, Jiajun Chen¹*, Lin Wang¹*

Department of Neurology, China-Japan Union Hospital of Jilin University, Changchun, China

* These authors contributed equally to this work.

* cjj@jlu.edu.cn (JC); wlmumu@jlu.edu.cn (LW)

Abstract

Introduction

Parkinson's disease (PD) is the second most common worldwide age-related neurodegenerative disorder without effective treatments. Cuproptosis is a newly proposed conception of cell death extensively studied in oncological diseases. Currently, whether cuproptosis contributes to PD remains largely unclear.

Methods

The dataset GSE22491 was studied as the training dataset, and GSE100054 was the validation dataset. According to the expression levels of cuproptosis-related genes (CRGs) and differentially expressed genes (DEGs) between PD patients and normal samples, we obtained the differentially expressed CRGs. The protein-protein interaction (PPI) network was achieved through the Search Tool for the Retrieval of Interacting Genes. Meanwhile, the disease-associated module genes were screened from the weighted gene co-expression network analysis (WGCNA). Afterward, the intersection genes of WGCNA and PPI were obtained and enriched using the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG). Subsequently, the key genes were identified from the datasets. The receiver operating characteristic curves were plotted and a PPI network was constructed, and the PD-related miRNAs and key genes-related miRNAs were intersected and enriched. Finally, the 2 hub genes were verified via qRT-PCR in the cell model of the PD and the control group.

Results

525 DEGs in the dataset GSE22491 were identified, including 128 upregulated genes and 397 downregulated genes. Based on the PPI network, 41 genes were obtained. Additionally, the dataset was integrated into 34 modules by WGCNA. 36 intersection genes found from WGCNA and PPI were significantly abundant in 7 pathways. The expression levels of the genes were validated, and 2 key genes were obtained, namely peptidase inhibitor 3 (PI3) and neuroserpin family I member 1 (SERPINI1). PD-related miRNAs and key genes-related

OPEN ACCESS

Citation: Song J, Li J, Pei X, Chen J, Wang L (2024) Identification of cuproptosis-related key genes and pathways in Parkinson's disease via bioinformatics analysis. PLoS ONE 19(4): e0299898. <https://doi.org/10.1371/journal.pone.0299898>

Editor: Qi Zhao, University of Science and Technology Liaoning, CHINA

Received: November 8, 2023

Accepted: February 17, 2024

Published: April 16, 2024

Copyright: © 2024 Song et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The datasets generated and analysed during the current study are available in the GEO (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE22491>, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE100054>).

Funding: This research was supported by Norman Bethune Program of Jilin University, Grant/Award Number: 2022B28; and the project "Development of human mesenchymal stem cell exosomes loaded with astaxanthin preparation and its

preclinical study for the treatment of Parkinson's disease" granted by the Jilin Provincial Science and Technology Department. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Abbreviations: AD, Alzheimer's disease; AUC, Area Under Curve; BP, biological process; CC, cellular component; CRG, cuproptosis-related gene; DEG, differentially expressed gene; ER, endoplasmic reticulum; FDR, False Discovery Rate; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MF, molecular function; PD, Parkinson's disease; PI3, peptidase inhibitor 3; PPI, protein-protein interaction; ROC, receiver operating characteristic; SERPIN1, neuroserpin family I member 1; STRING, Search Tool for the Retrieval of Interacting Genes; TCA, tricarboxylic acid; WGCNA, weighted gene co-expression network analysis.

miRNAs were intersected into 29 miRNAs including hsa-miR-30c-2-3p. At last, the qRT-PCR results of 2 hub genes showed that the expressions of mRNA were up-regulated in PD.

Conclusion

Taken together, this study demonstrates the coordination of cuproptosis in PD. The key genes and miRNAs offer novel perspectives in the pathogenesis and molecular targeting treatment for PD.

1. Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disease with a global prevalence of more than 6 million individuals [1], second only to Alzheimer's disease (AD). To date, there are no agents shown to have unequivocal evidence of disease-modifying effects in PD, which is hazardous to human health [2]. Common clinical manifestations include motor symptoms such as tremor, myotonia, bradykinesia, and postural balance disorders, as well as non-motor symptoms such as sensory and autonomic symptoms [3]. Pathologically, PD is characterized by a loss of dopaminergic neurons in the substantia nigra and the presence of Lewy bodies composed of abnormal accumulation of α -synuclein in the midbrain [4]. However, the pathogenesis of PD is not conclusive and the metal ion dyshomeostasis [5], oxidative stress [6], mitochondrial dysfunction [7], and chronic inflammation [8] may be involved. In addition, genetic factors play an important role in the onset and progression of PD. Generally, rare DNA variants are typically associated with monogenic or familial PD; and more common, smaller effect variants are usually identified in apparently sporadic PD [9]. Therefore, identifying related genes and exploring the possible pathways that may be involved in PD pathogenesis could provide new ideas for the prediction and intervention of the disease.

Copper is important for various neuronal functions such as interacting with synaptic proteins and neurotransmitter receptors in synapses [10]. In vitro studies have shown that the presence of copper in millimolar concentrations causes the formation of partially folded amyloidogenic conformations with a propensity for aggregation [11]. Mechanically, copper dysfunction is associated with different toxic effects, mainly represented by oxidative stress [12]. Therefore, careful homeostatic control of copper levels plays an important role in PD pathology, leading to the development of novel therapeutic approaches based on restoring copper homeostasis [11]. Currently, experimental evidence has proven the inhibition of α -synuclein oligomer-mediated reactive oxygen species by copper chelators [13]. Cuproptosis is a novel form of cell death newly proposed [14]. It occurs via the direct binding of copper to lipoylated components of the tricarboxylic acid (TCA) cycle in the mitochondria, leading to proteotoxic stress and ultimately cell death. As is well acknowledged, the key role of mitochondrial dysfunction has been demonstrated in PD, and several potential therapeutic avenues targeting to promote the clearance of old or damaged mitochondria have been consistently developed [15]. A recent study showed that isogenic human induced pluripotent stem cells with PARK2 knockout displayed abnormal TCA cycle activity, perturbed mitochondrial ultrastructure, and increased oxidative stress [16]. Therefore, it would be reasonable to infer that cuproptosis is closely associated with PD. However, the potential regulatory mechanisms remain unknown and need further exploration. What's more, the identification of key genes and miRNAs in PD might assist to explain the correlation from a genetic perspective.

Predicting key differentially expressed genes (DEGs) for disease by comprehensive biochemical computational methods is a prevalent approach that efficiently and rapidly offers molecular insights into the disease utilizing a variety of resources of sequencing results in a high-throughput manner. Currently, there have been some studies predicting the cuproptosis-related biomarkers of the neurodegenerative diseases based on bioinformatics tools [17–20]. Lai et al. [17] combined the weighted gene co-expression network analysis (WGCNA) algorithm with several machine models such as random forest model to explore the cuproptosis-related molecular clusters and construct a prediction model for AD. As for PD, few studies [18–20] aimed to discover the DEGs about cuproptosis by means of different computational biological methods. A recent bioinformatics analysis [20] found that the cuproptosis-related key genes intervened in the progression of PD through the integration of DEGs, WGCNA-related significant module genes (with the soft threshold of 10) and immune cell infiltration, resulting in three key genes different from our study, which might be attributed to the different selection of parameters. However, all the genes obtained from the comprehensive bioinformatics analyses would be of great importance for the exploration of molecular perspectives on disease.

In this research, we acquired 10 PD and 8 normal samples in the GSE22491 database as the training dataset, and 10 PD and 9 normal samples in the GSE100054 as the validation dataset. Based on the expression levels of cuproptosis-related genes (CRGs) and DEGs between PD and normal samples, a correlation analysis was performed to observe the differentially expressed CRGs. A protein-protein interaction (PPI) network was performed using the Search Tool for the Retrieval of Interacting Genes (STRING). Meanwhile, we screened for disease-associated module genes by the WGCNA and took the intersection of significantly related modular genes in the results of WGCNA and the genes from PPI, which were enriched by the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) in succession. Then the expression levels of the intersection genes were validated in the training and validation datasets to identify two key genes, peptidase inhibitor 3 (PI3) and neuroserpin family I member 1 (SERPINI1). Next, we plotted the receiver operating characteristic (ROC) curves in all datasets and constructed a PPI network. Finally, 29 key miRNAs were screened by taking the intersection of PD-related miRNAs and key genes-related miRNAs and enriched by the KEGG analysis. A detailed flow chart of the bioinformatics analysis was exhibited (Fig 1). Finally, the mRNA expressions of PI3 and SERPINI1 were verified from the qRT-PCR results. Through the above research, it is concluded that cuproptosis has effects on PD pathogenesis. Moreover, these genes and factors may be potential targets for therapeutic drugs.

2. Materials and methods

2.1 Data acquisition

The datasets used in this analysis were obtained from the NCBI GEO database [21] (<https://www.ncbi.nlm.nih.gov/>). We searched for PD-related datasets using "Parkinson's disease", obtained the pre-processed, normalized and log₂-transformed probe expression matrix, followed by downloading the platform annotation file. The probes that did not match the Gene symbol were removed. For different probes mapping to the same gene, the mean value of different probes was taken as the final expression value of the gene. GSE22491 was selected as the training dataset with 18 samples, including 10 PD and 8 normal samples, and the sequencing platform was GPL6480 (Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)). In addition, GSE100054 was selected as the validation dataset. 19 samples, including 10 PD and 9 normal samples, were sequenced on GPL23126 ([Clariom_D_Human] Affymetrix Human Clariom D Assay [transcript (gene) version]). Data were freely

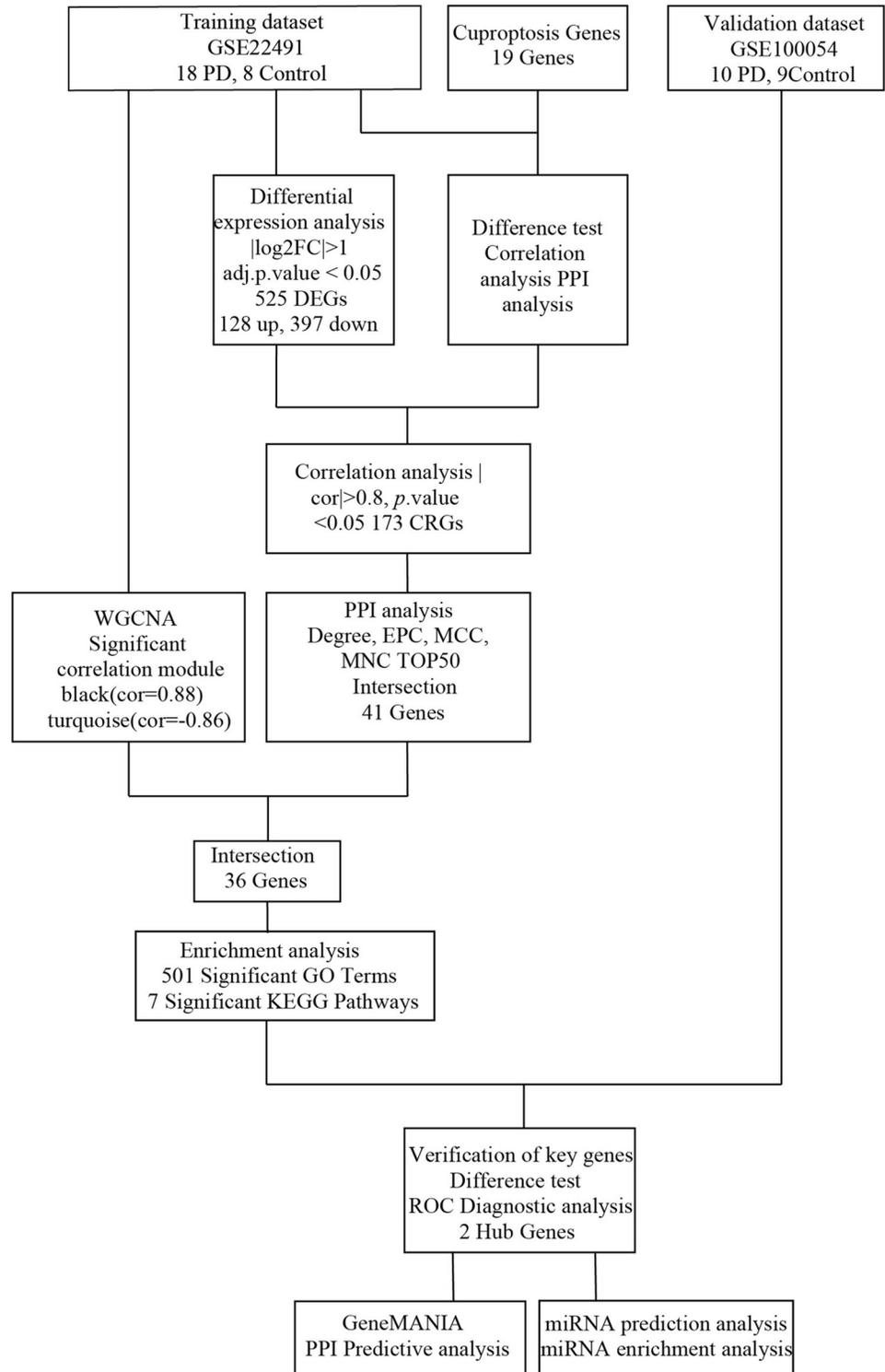


Fig 1. The flow chart of the bioinformatics analysis.

<https://doi.org/10.1371/journal.pone.0299898.g001>

available online, and our study did not involve any experiments in the lab performed by any of the authors.

2.2 Differential analysis of CRGs

19 CRGs were obtained from the relevant literature [22]. The expression levels of the genes were extracted for differential analysis from the training dataset between the PD and normal samples. The rank sum test was used for differential analysis. It is the method of hypothesis testing commonly used between two groups of samples to test whether the difference of a certain indicator is significant. Then a correlation analysis was performed among the genes. Correlation analysis was performed using the pearson algorithm of cor function in R (version 3.6.2). Pearson correlation analysis is a method to analyze the linear correlation between the two variables. The PPI network was analyzed using the STRING [23] (Version: 10.0, <http://www.string-db.org/>) database, with the input gene set of CRGs and species of homo. A score of 0.4 (medium confidence) was selected as the cutoff criterion, requiring that the interacting protein nodes were included in the above genes.

2.3 DEGs and differentially expressed CRGs screening

The R software limma package [24] (Version 3.10.3, <http://www.bioconductor.org/packages/2.9/bioc/html/limma.html>) was used to perform DEGs analysis in the training dataset. Limma, a type of generalized linear model that fits a linear equation to the expression of each gene, is a common method for screening for the DEGs. False Discovery Rate (FDR) < 0.05 and $|\log_{2}FC| > 1$ were set as the threshold to screen the DEGs.

Correlation analysis of CRGs and DEGs was performed using the pearson algorithm of cor package in R. DEGs whose correlation coefficients with at least one CRG were greater than 0.8 were selected as differentially expressed CRGs, which were used for further screening analysis.

2.4 PPI network construction for differentially expressed CRGs

The reciprocal relationships among differentially expressed CRGs encoding proteins were analyzed from the STRING database. The database contains 1380838440 interactions of 9643763 proteins from 2031 species, which could be used for protein interaction analysis. The graph was constructed using the Cytoscape software [25] (version 3.4.0, <http://chianti.ucsd.edu/cytoscape-3.4.0/>).

In addition, we applied the CytoNCA [26] plug-in (Version 2.1.6, <http://apps.cytoscape.org/apps/cytonca>) to analyze the nodal topological properties of the network, where Degree, EPC, MCC, and MNC were used as the main attributes. CytoNCA is a cytoscape plugin for centralized analysis and evaluation of biological networks, so that the key nodes could be filtered from the interworking networks. The larger value of each attribute proved the greater role of the gene in the network. The TOP 50 genes were selected under each attribute in turn and the key differentially expressed CRGs were obtained by taking the intersection.

2.5 WGCNA screening for disease-related modules

WGCNA is a systematic biological method to characterize the modules of gene association between samples. It could be used to identify the highly synergistic gene modules which could be served as the candidate biomarkers or therapeutic targets based on the introgression and association between the gene set modules and the disease traits. The input genes were analyzed using the R package WGCNA [27] (Version 1.61, <https://cran.r-project.org/web/packages/WGCNA/>) based on the dataset GSE22491 to define modules of highly correlated genes

associated with PD. In the WGCNA algorithm, the element in the defined gene co-expression modules was the weighted value of the correlation coefficient, and it was chosen so that the networks were scale-free. The weighted value here was the soft Power. The parameters (min Module Size = 30; merge Cut Height = 0.25) were set to aggregate highly correlated genes into modules based on the clustering and dynamic pruning methods. Finally, the modules that are correlated with the disease with the P -value less than 0.05 were selected as disease-related modules by analyzing the correlation between the module and the clinical traits (whether the sample is a diseased or normal sample). The genes in the modules with the most positive and negative correlation were selected as the genes associated with PD.

2.6 Intersection genes screening

Disease-related module genes were obtained by WGCNA, and the key differentially expressed CRGs were obtained by the PPI analysis. They were intersected for further study.

2.7 Gene enrichment analysis

The enrichment analysis is a statistical method widely used in the bioinformatics analyses, which is mainly applied to identify the biologically significant patterns and functions from a large amount of genetic information. GO analysis is commonly used to test the enrichment of Gene Ontology entries in a collection of genes and predict the common features of the genes in terms of biological process (BP), molecular function (MF) and cellular component (CC), while KEGG is used to analyze the role of genes in metabolic and signaling pathways.

The GO [28] function enrichment analysis and KEGG [29] pathway enrichment analysis were performed using R package clusterProfiler [30]. The GO analysis terms included CC, MF, and BP. $P < 0.05$ was considered significant due to the few enriched pathways.

2.8 Identification of key genes

The box plots of the screened genes were used to show the expression levels in the training and validation datasets. The rank sum test was used to identify the key genes ($P < 0.05$), requiring they were consistent up or down-regulation trend in the datasets.

2.9 Diagnostic accuracy assessment of key genes and PPI network construction

The pROC package Version 1.12.1 [31] (<https://cran.r-project.org/web/packages/pROC/index.html>) in R language was used to plot ROC curves in all datasets to assess the diagnostic accuracy. The Area Under Curve (AUC) was utilized to assess the area under the ROC curve enclosed with the coordinate axis. The PPI analysis of the key genes was performed using the GeneMANIA database (<https://genemania.org/>) [32]. The database could predict the proteins in terms of co-localization, shared protein structural domains, and the correlation with the signaling pathways, etc.

2.10 Key miRNAs screening

HMDD is a hand-collected database to explore the miRNAs associated with the disease. The HMDD V3.0 database [33] (<http://www.cuilab.cn/hmdd>) was used to retrieve PD-related miRNAs, and miRWalk (<http://mirwalk.umm.uni-heidelberg.de/>) [34] was used to predict miRNAs for key genes. They were intersected to obtain the key miRNAs, and a gene-miRNA network was constructed.

2.11 KEGG pathway enrichment analysis of key miRNAs

The KEGG pathway analysis was performed on the key miRNAs through DIANA-miRPath v3.0 [35] (<http://www.microrna.gr/miRPathv3/>). DIANA-miRPath is a web server for miRNA pathway analysis, through which the miRNA enrichment could be analyzed. Pathways showing $FDR < 0.05$ were considered significantly enriched.

2.12 RT-qPCR confirmation

Reverse transcription qPCR was used to quantify the amount of mRNA measured by the fluorescence quantitative PCR instrument from Beckman, USA. The cell model of PD was constructed with 1 mM MPP+ (sigma) treatment of SH-SY5Y cells. Total RNA from SH-SY5Y cells was extracted with Trizol reagent (leagene, China). The cDNA was generated immediately from 1 μ g extracted RNA by HiScript II Q RT SuperMix for qPCR (Vazyme, China). Then quantitative PCR was performed by ChamQ Universal SYBR qPCR Master Mix (Vazyme, China). The relative gene expression was analyzed by means of the $2^{-\Delta\Delta C_t}$ method. The primers were designed to target PI3 and SERPINI1. The sequences are as follows: PI3: Forward: TGTCAAAGGCCGTGTTCCAT, Reverse: GAGCCAGGCTTAGTGAGAC; SERPINI1: Forward: AGCAATTCAAGTCCTTCCATAGAG, Reverse: CTTGAGGATACAGCACAGCCATC.

3. Results

3.1 Differential analysis of CRGs

Based on the expression levels of CRGs (S1 Table) in the training dataset, differential analysis between PD and normal samples was performed and a box plot was drawn (Fig 2A). To explore the association of CRGs, the Pearson's correlation coefficient was calculated and a heatmap was drawn to present the results (Fig 2B). The PPI network of CRGs was performed based on the STRING database (Fig 2C).

3.2 Differentially expressed CRGs screening and PPI network construction

128 up-regulated and 397 down-regulated genes were obtained from the training dataset, altogether 525 DEGs. Based on the screened differential genes, a volcano plot was performed (Fig 3A).

Correlation analysis of CRGs and DEGs was performed and 173 differentially CRGs were screened in total. Further, a PPI network was performed (Fig 3B). We obtained a reciprocal network consisting of 79 genes, demonstrating that these genes had close interactions and might play an important role in the disease. The nodal topological properties of the network were analyzed, and the TOP 50 genes of the four attributes were selected to take the intersection (Fig 3C). 41 key differentially expressed CRGs were observed.

3.3 WGCNA and intersection genes screening

WGCNA was performed on the training dataset, and the soft threshold was 8 (Fig 4A and 4B). Next, the highly correlated genes were aggregated into modules based on the clustering and dynamic pruning method, and these modules were clustered with the coefficients of dissimilarity of 0.25. Highly correlated genes were aggregated into 34 modules. Further, the correlation between the eigenvector gene of each module and the clinical traits was calculated and the modules with a significant *P*-value was plotted (Fig 4C). The black module (540 genes; correlation coefficient $r = 0.88$ and $P.value < 0.001$) showed the strongest positive correlation with PD, so this module was taken as the key module associated with PD. The turquoise module (5916

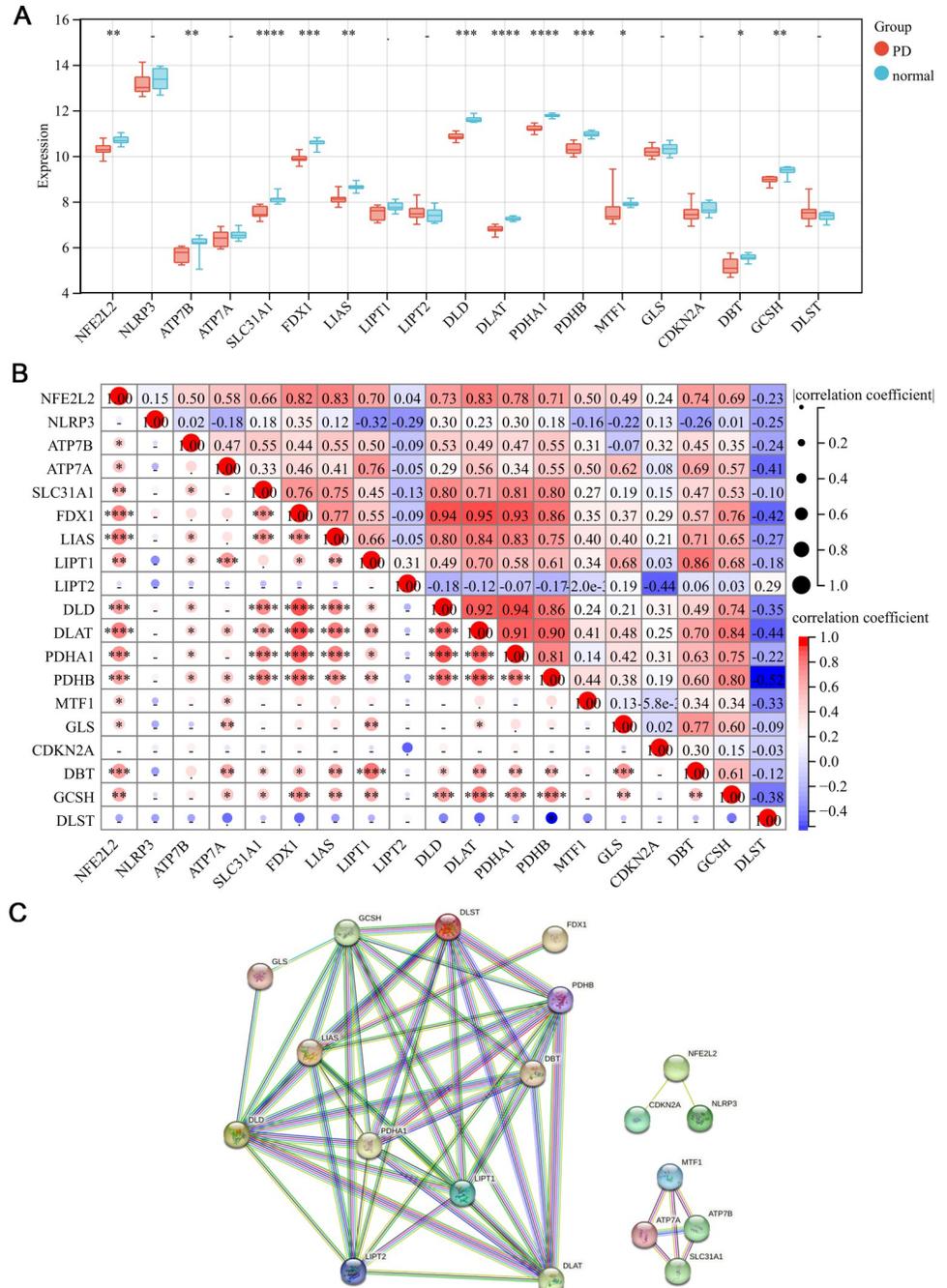


Fig 2. The differential analysis of CRGs. (A) The box plot of CRGs expression levels between PD and normal samples. (B) The heatmap of CRGs. (C) The PPI network of CRGs. *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$; ****: $P < 0.0001$; the remaining symbols: not statistically significant ($P \geq 0.05$).

<https://doi.org/10.1371/journal.pone.0299898.g002>

genes; correlation coefficient $r = -0.86$ and P .value < 0.001) showed the strongest negative correlation with PD, so this module was also considered a PD-related key module.

There were 36 genes (S2 Table) obtained by intersecting the black and turquoise modular genes from WGCNA and the genes from PPI (Fig 4D).

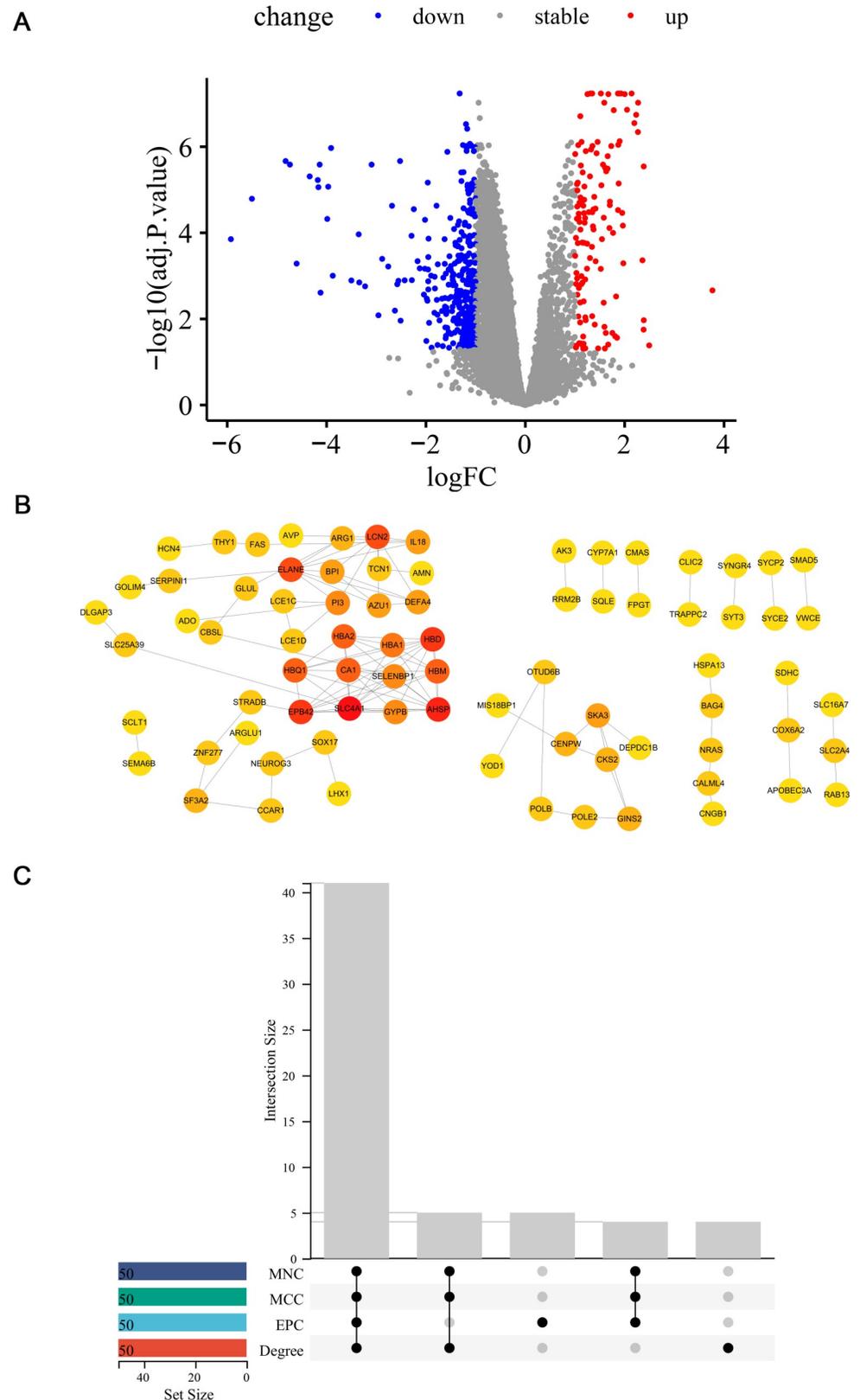


Fig 3. DEGs screening and PPI network construction. (A) The volcano plot of DEGs (red indicated up-regulated genes, blue indicated down-regulated genes, and gray indicated genes with insignificant differences of expression

levels). (B) The PPI network construction for differentially CRGs. The gray connecting lines indicated the interactions between the corresponding proteins of genes. Red indicated genes with more connecting lines, and yellow indicated genes with fewer connecting lines. (C) The upset diagram taking the intersection of the TOP 50 genes of each nodal topological property in the PPI network.

<https://doi.org/10.1371/journal.pone.0299898.g003>

3.4 Gene enrichment analysis

GO and KEGG were performed on the above-obtained genes. 501 terms were enriched by the GO analysis, including oxygen transport, gas transport, cellular response to toxic substance, specific granule lumen, cytoplasmic vesicle lumen, haptoglobin binding, hemoglobin binding, iron ion binding, etc. The key terms were selected for graphical display (Fig 5A). 7 significantly enriched pathways (S3 Table) were obtained from the KEGG analysis (Fig 5B). The specific signaling pathways included African trypanosomiasis, malaria, nitrogen metabolism, arginine biosynthesis, base excision repair, biosynthesis of amino acids, and sulfur metabolism.

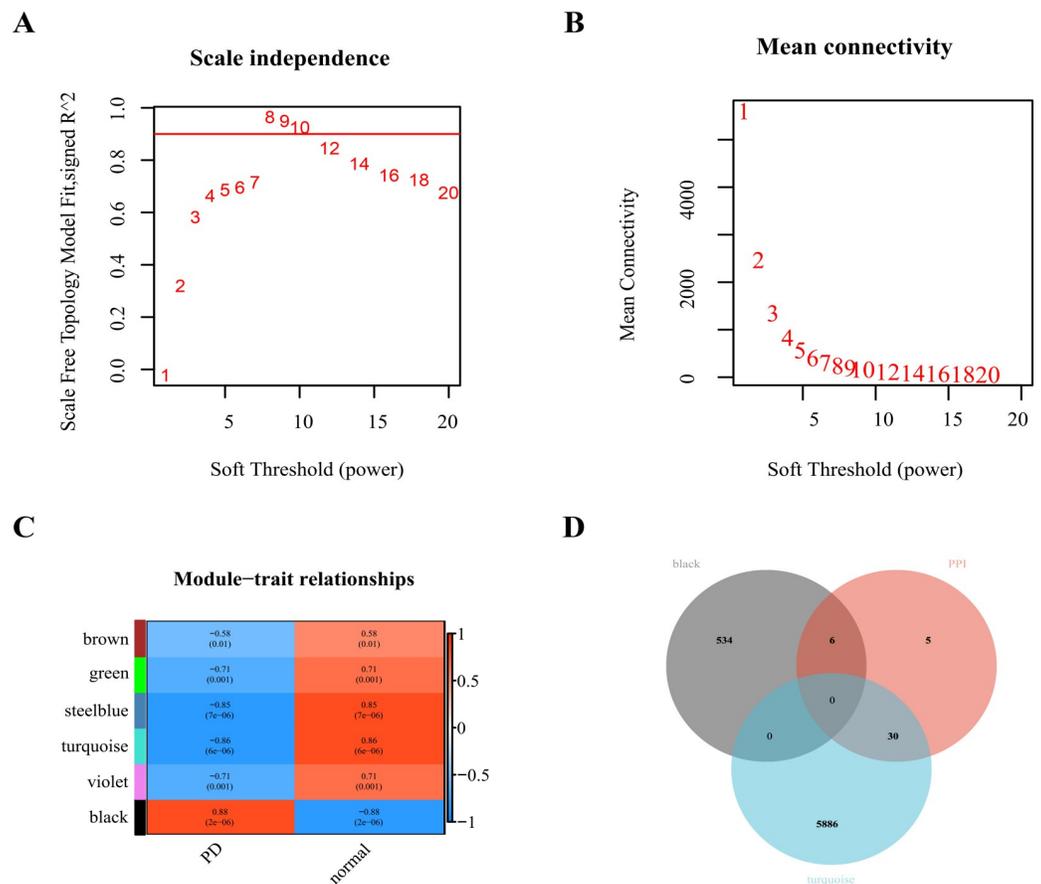


Fig 4. WGCNA and intersection genes screening. (A) The analysis of network topology for various soft-thresholding powers. The x-axis reflected the soft-thresholding power, and the y-axis reflected the scale-free topology model fit index. Here we selected the soft threshold corresponding to the ordinate reaching 0.9 for the first time, which was 8. (B) The x-axis reflected the soft-thresholding power, and the y-axis reflected the mean connectivity (degree). (C) The correlation analysis between gene modules and clinical traits. The upper numbers indicated the correlation coefficients, and the lower bracketed numbers indicated P values. (D) The Venn diagram obtained by taking the intersection of PD-related key module genes and PPI key genes.

<https://doi.org/10.1371/journal.pone.0299898.g004>

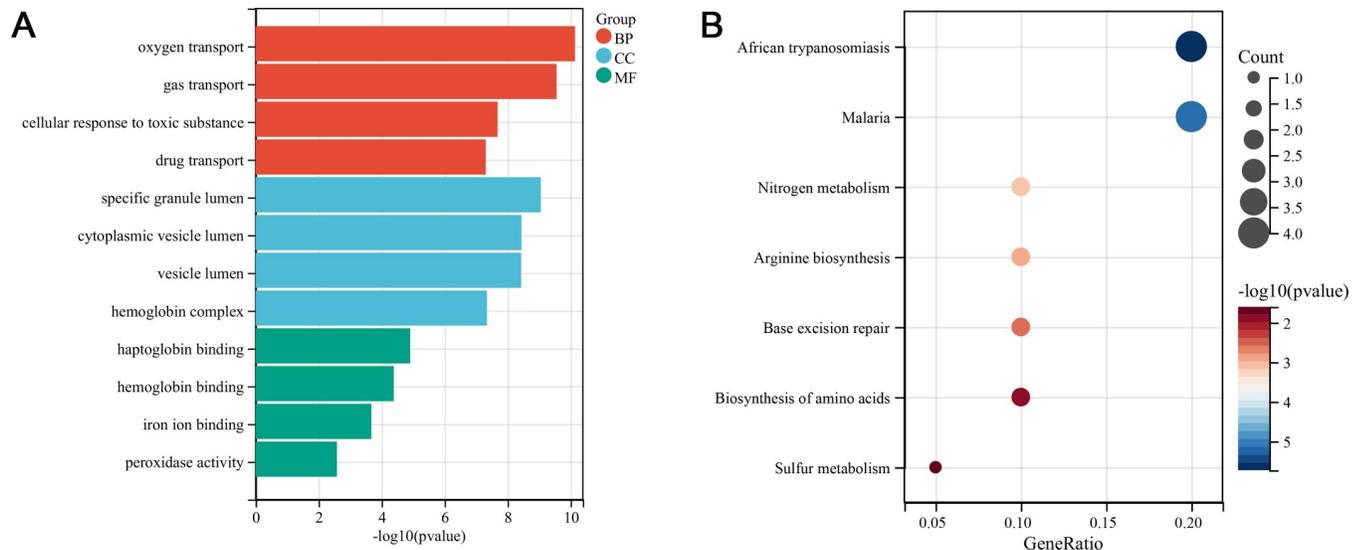


Fig 5. Gene enrichment analysis. (A) The GO term histogram of the intersection genes. The 4 GO terms with the smallest P values in each group was displayed. (B) The KEGG pathway bubble graph of the intersection genes.

<https://doi.org/10.1371/journal.pone.0299898.g005>

3.5 Identification, diagnostic accuracy assessment, and PPI network construction of key genes

6 significantly different genes, including SERPINI1, PI3, glutamate-ammonia ligase, late cornified envelope 1C, lipocalin 2, and interleukin 18 (S4 Table), were found between PD and normal samples in the validation dataset, of which 2 down-regulated genes, PI3 and SERPINI1, were consistent in the two datasets. The box plots were drawn for the expression levels of the two genes in the training and validation datasets (Fig 6A–6D).

The ROC curves were plotted in all datasets of the key genes to assess the diagnostic accuracy through the pROC package in R language. It turned out that the AUC values for the key genes were greater than 0.7 in both datasets (Fig 7A–7D). What's more, the PPI analysis was performed using the GeneMANIA database (Fig 8). The main significantly enrichment pathways included endopeptidase inhibitor activity, peptidase inhibitor activity, endopeptidase regulator activity, peptidase regulator activity, enzyme inhibitor activity, antimicrobial humoral response and humoral immune response, etc.

3.6 Key miRNAs screening and KEGG pathway enrichment analysis

Based on the HMDD database, 131 miRNAs related to PD were obtained and intersected with the miRNAs associated with key genes. The miRNAs with a score of 1 were selected to obtain 29 key miRNAs like hsa-miR-30c-2-3p, hsa-miR-34a-5p, hsa-miR-4697-5p, etc. A network was constructed based on the key genes and miRNAs (Fig 9A), which were enriched by KEGG (Fig 9B). The main KEGG pathways encompassed cell cycle, proteoglycans in cancer, protein processing in endoplasmic reticulum, adherens junctions and chronic myeloid leukemia.

3.7 RT-qPCR confirmation

The expression of 2 hub genes, PI3 and SERPINI1 were verified by the integration bioinformatics analysis. Based on the qRT-PCR results, the expressions of PI3 ($P = 0.045$, Fig 10A) and SERPINI1 ($P = 0.008$, Fig 10B) were up-regulated in PD compared to the control group, which were inconsistent with the results of the bioinformatics analysis.

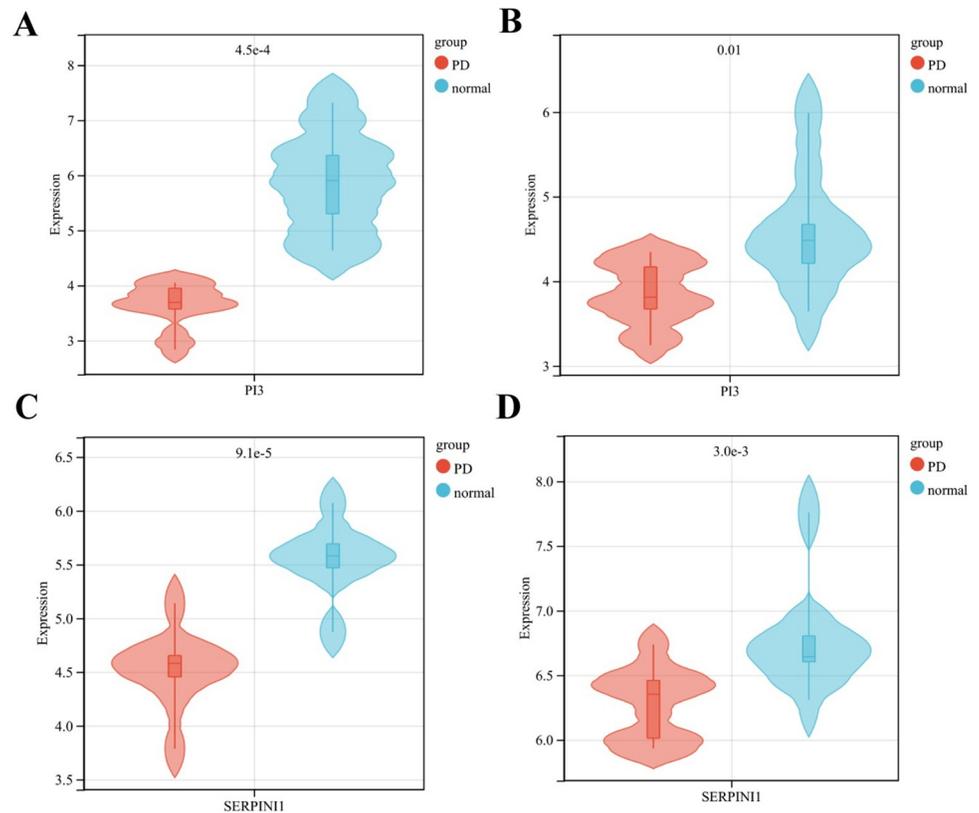


Fig 6. Identification and diagnostic accuracy assessment of key genes. The box plots of PI3 expression levels between PD and normal samples in the training dataset (A) and validation dataset (B), and SERPINI1 expression levels between PD and normal samples in the training dataset (C) and validation dataset (D).

<https://doi.org/10.1371/journal.pone.0299898.g006>

4. Discussion

PD is a kind of disease that seriously endangers human health globally. Although there is rapid progress and evolution in technology to unravel the pathogenic factors of PD, the underlying pathogenesis remains to be determined. Copper is a transition metal linked to pathological and beneficial effects in neurodegenerative diseases [36], as both excess copper levels and copper deficiency can be harmful [11]. There is evidence that excess copper leads to neuronal cell death and α -synuclein aggregation [36, 37]. Accordingly, Cu^{II} [atsm] has been demonstrated to rescue dopaminergic cell loss and improve motor dysfunction based on both sporadic (toxin-based) and genetic mouse models of PD [38]. A meta-analysis based on the 18 eligible studies identified that substantia nigra copper levels in PD patients were significantly lower than the control cases [39]. Consequently, the controversial association should be further studied to reveal the underlying mechanisms. Cuproptosis is a mode of cell death associated with lipoylated TCA enzymes in mitochondrial metabolism, different from apoptosis, ferroptosis, and necroptosis [14]. It has attracted tremendous interests in the field of tumor studies. However, it needs further investigation to determine the biological processes and signal pathways related to cuproptosis in PD.

In this study, we acquired PD and normal samples from the database as the training and validation datasets and compared the expression levels of CRGs and DEGs between PD and normal samples to obtain the differentially expressed CRGs. The disease module-related genes and the key differentially expressed CRGs were respectively obtained by WGCNA and PPI,

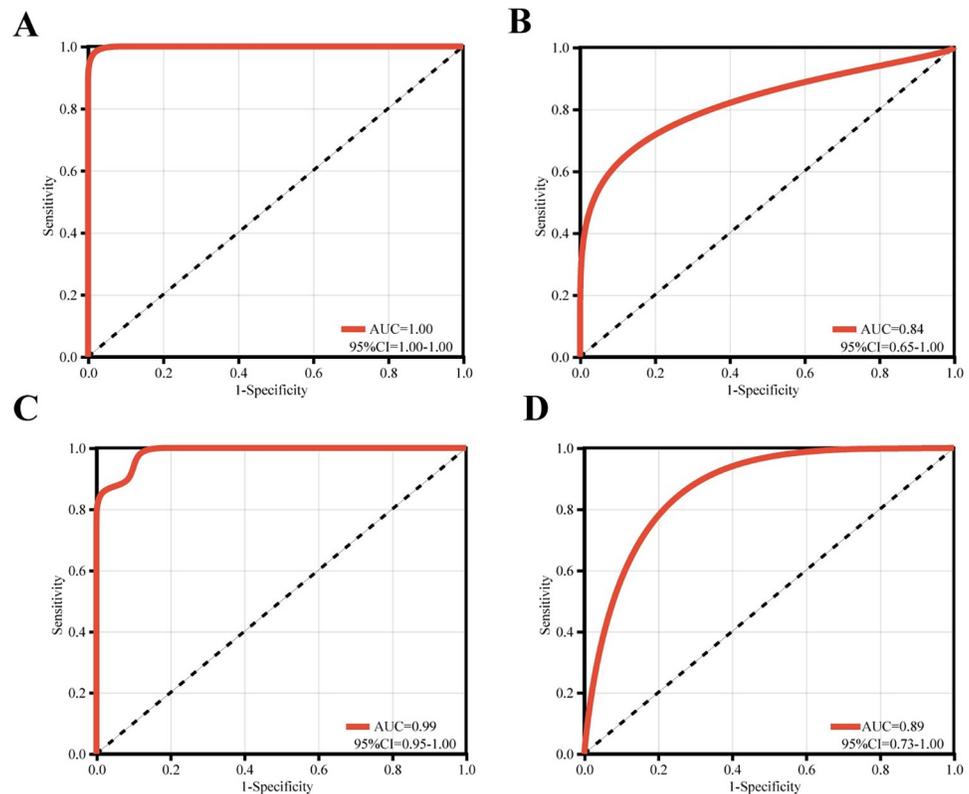


Fig 7. The ROC curves of PI3 in the training dataset (A) and validation dataset (B), and that of SERPINI1 in the training dataset (C) and validation dataset (D).

<https://doi.org/10.1371/journal.pone.0299898.g007>

and they were intersected to observe the intersection genes. Then we enriched the genes by the KEGG and GO analysis. Next, we identified 2 key genes from the datasets, PI3 and SERPINI1, and constructed a PPI network. They were verified by the ROC curves ($AUC > 0.7$). Then the PD-related miRNAs and key genes-related miRNAs were intersected into key miRNAs, which were enriched by the KEGG analysis. At last, the results were verified by RT-qPCR. Technically, we integrated the bioinformatics analysis and the biological experiment to figure out not whether cuproptosis-related genes were involved in the pathogenesis of PD. To date, we utilized these methods for the first time to explore the association between cuproptosis and the disease, innovatively elucidating the underlying pathogenesis from a new perspective and laying the foundation for subsequent basic experiments. It's worth mentioning that the intervention against the corresponding cuproptosis-related molecules would provide promising targets for PD therapy as well as diagnostic biomarkers deserving further investigation.

From the data analysis in this study, we obtained 2 key genes, PI3 and SERPINI1, respectively. PI3, which functions in preventing excessive tissue injury during inflammatory events [40], has a significant association with a diversity of tumors [41]. Besides, one recent study based on pathway clustering analysis found that PI3 is upregulated in PD patients [42], which was consistent with our experimental results. Therefore, how PI3 participates in PD patients could be investigated in the future to explain this contradictory finding, possibly related to the different stages of neuroinflammation. SERPINI1 encodes an axonally secreted neuroprotective protein called neuroserpin, which values in synaptic plasticity [43]. It is implicated in the conformational disease familial encephalopathy with inclusion bodies, which manifested as progressive dementia and epilepsy via endoplasmic reticulum (ER)-overload response [44].

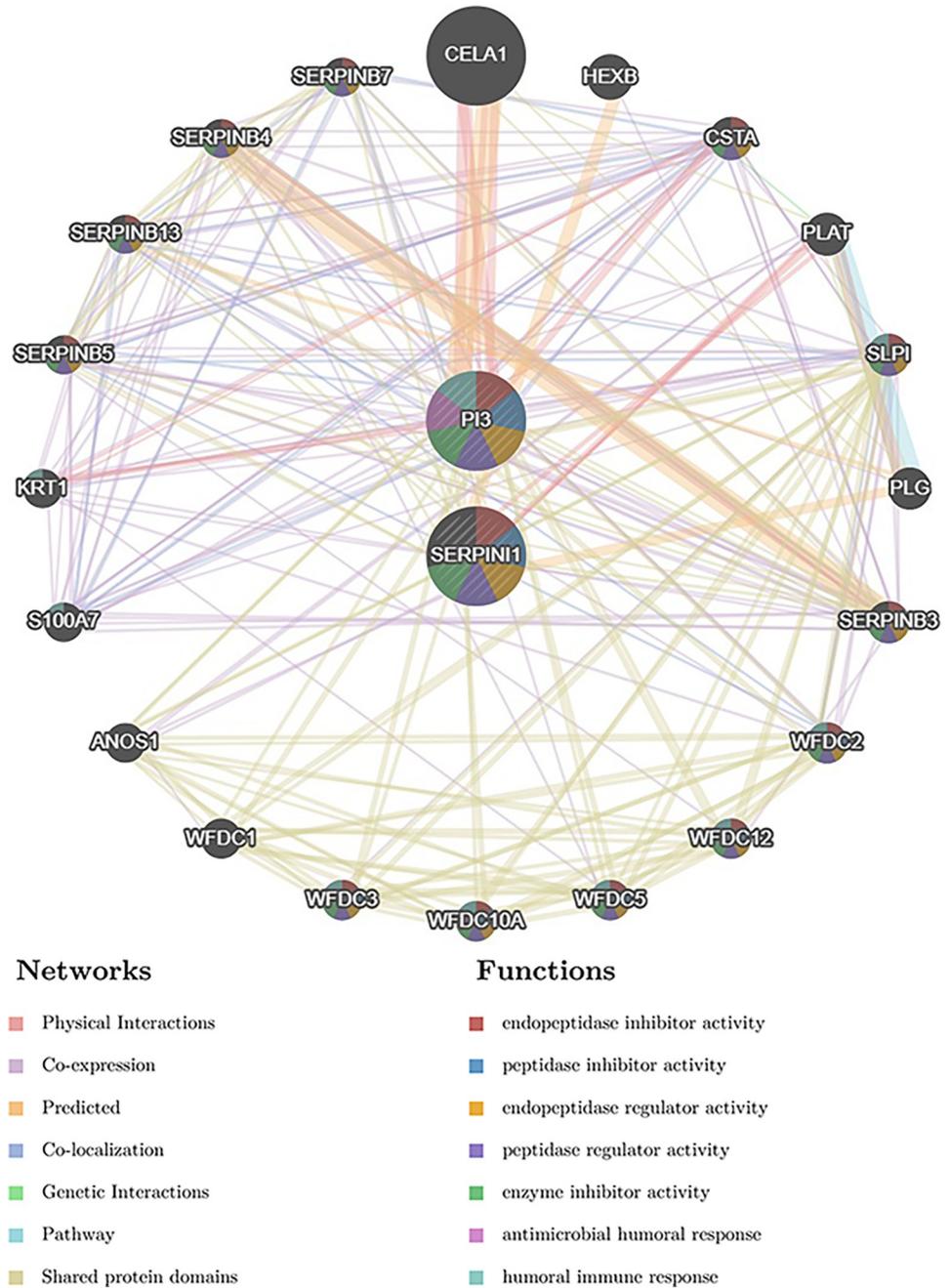


Fig 8. The PPI network of the key genes.

<https://doi.org/10.1371/journal.pone.0299898.g008>

In a study based on analyzing human SERPIN transcripts, SERPINI1 was significantly down-regulated in Sporadic Creutzfeldt-Jakob disease patients [45]. Besides, the role of SERPINI1 in AD pathology is controversial. Neuroserpin can act neuroprotective by binding to Aβ and altering its oligomerization [46], however, it could also be detrimental in reducing the clearance of Aβ [47]. The conflicting role in AD is similar to the gap between our bioinformatics predictions and experimental results, which requires further studies. What's more, aggregated α-synuclein has been demonstrated to induce ER fragmentation and compromise ER protein

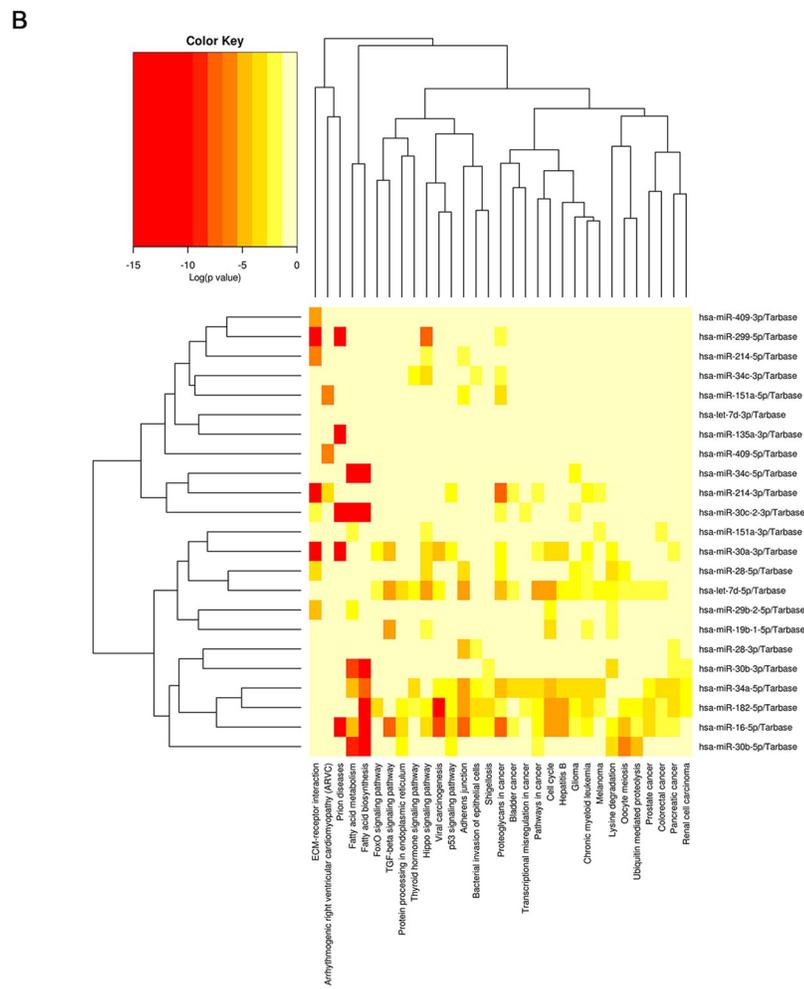
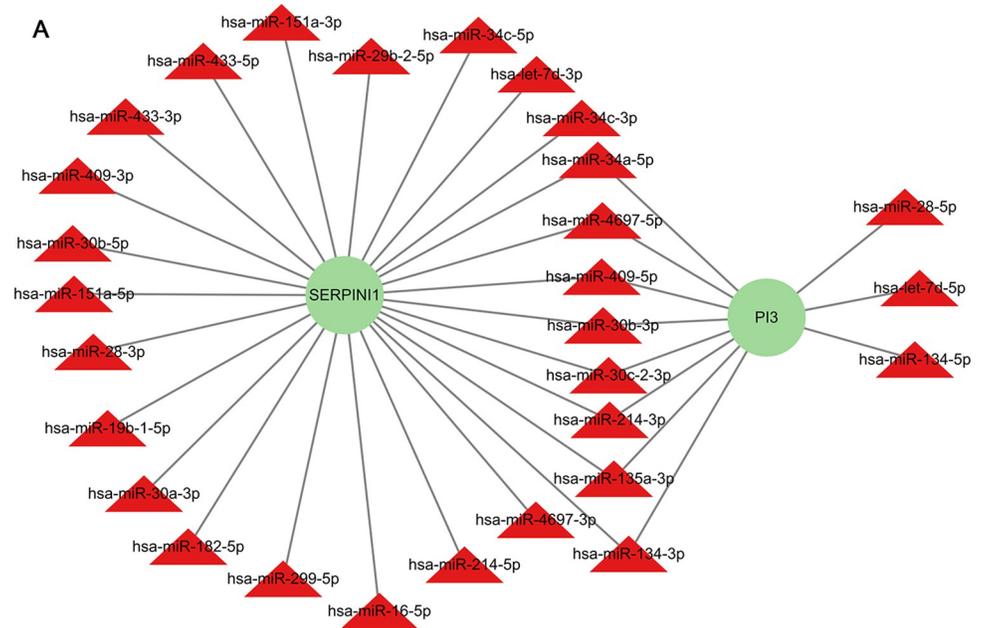


Fig 9. Screening and KEGG pathway enrichment analysis of key miRNAs. (A) The network of key genes and miRNAs (the red triangles were key miRNAs and the green circles were key genes). (B) The significantly enriched KEGG pathways of key miRNAs (red indicated more significantly enriched).

<https://doi.org/10.1371/journal.pone.0299898.g009>

folding capacity [48]. Therefore, the role of SERPINI1 could be further investigated in other neurodegenerative diseases like PD targeting α -synuclein and ER homeostasis. Chymotrypsin-like elastase family member 1 was predicted as a key interacting gene by PPI analysis of the key genes. It is a digestive protease expressed during lung development and regeneration [49]. However, its role in PD remains unclear and needs to be further clarified.

What's more, the interaction prediction models targeting gene, mRNA, miRNA, lncRNA [50, 51] proteins [52], metabolites [53, 54] and even the drug compounds [55, 56] have developed in the fields of computational biology, which could provide fresh and valuable insights into the genetic markers related with the disease. In recent years, with the popularity of single-cell sequencing, the deep-learning methods for single-cell clustering analysis have also emerged consequentially with great prospects for development [52, 57]. Generally, the rapid growth of the computational biological methods results in a deep and comprehensive insight of the disease from the molecular level. Herein, by intersecting the related miRNAs of PD and key genes, we screened 29 key miRNAs including hsa-miR-30c-2-3p, hsa-miR-34c-5p, and hsa-miR-30b-5p, etc. Consistent with our results, another study found that the level of hsa-miR-30c-2-3p in plasma extracellular vesicles was significantly higher in PD patients than that in controls, and the target genes were enriched in dopaminergic synapse and PD pathway [58]. Eventually, we predicted fatty acid biosynthesis and fatty acid metabolism as the key pathways by KEGG. Considering the significance of lipoylated proteins in cuproptosis, it is reasonable to infer that cuproptosis takes part in the pathogenesis of PD.

Currently, there are few bioinformatic analyses targeting cuproptosis in PD as mentioned previously [18–20]. A recent study was dedicated to exploring the potential CRGs, the immune infiltration patterns in PD, and consequently, the associated copper chelators therapy [18]. Another study showed that CRGs were significantly enriched in immunity-related pathways like JAK-STAT and Notch pathway. Based on the LASSO analysis, core genes were identified

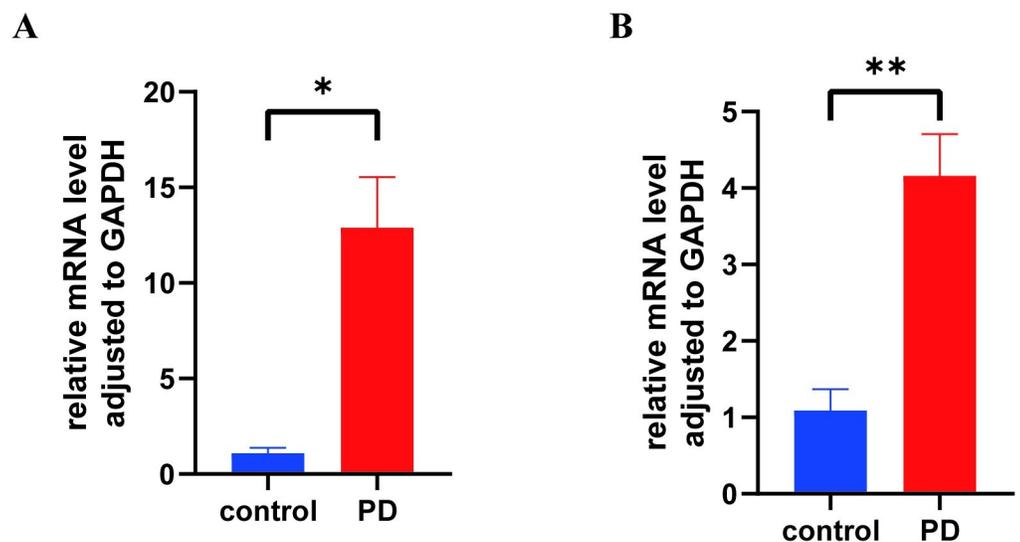


Fig 10. The RT-qPCR results of 2 hub genes. (A) The RT-qPCR results of PI3 between PD and the control group. (B) The RT-qPCR results of SERPINI1 between PD and the control group. *: P < 0.05; **: P < 0.01.

<https://doi.org/10.1371/journal.pone.0299898.g010>

and utilized in a prognostic prediction model [19]. Recently, Zhang et al. found that Cu exposed to mice lead to neuronal degeneration and promoted the expression of cuproptosis-related proteins ferredoxin 1 and dihydrolipoamide S-acetyltransferase, providing experimental evidence for the involvement of cuproptosis in neurodegenerative diseases [59]. In contrast to the above articles, the systematic bioinformatic and experimental methods were simultaneously used in our study to screen and validate the hub genes. Above all, the identification of key genes and miRNAs is of guiding value for exploring the underlying pathogenesis, the biomarkers for early identification, and the potential candidates for drug targeting.

However, it remains some limitations in our study as listed. 1) the limited sample size we obtained from public databases may not only affect the accuracy of predicting the key genes but also restrict the subgroup analysis, so it could be further designed in a larger size of the samples. What's more, the samples could be stratified by gender and other demographic characteristics, and PD could be categorized into different subtypes in future studies. 2) The algorithms in our study could be affected by the adoption of the coefficients and calculation methods, which could have impact on the results. Future research could integrate diverse computational biology algorithms and methods, such as ordinary differential equations based theoretical models on gene and protein signaling networks [60–62], to obtain the results robustly from different perspectives. 3) The results of the experimental validation are contradictory to our predictions. The complicated results might be related with the post-translation modification, the different stages of the disease, etc. Therefore, the in-depth studies could be conducted to figure out the underlying mechanisms.

5. Conclusion

In summary, our study demonstrates that cuproptosis contributes to the development of PD through comprehensive bioinformatics approaches combined with experimental verification. Future studies could be designed to figure out the pathogenic role of cuproptosis in PD based on larger populations and other algorithms like ordinary differential equations based theoretical models. The key genes and mi-RNAs obtained from the study could be instructive in the identification of the diagnostic biomarkers as well as the candidates for molecular modification and metal homeostatic therapies in the early course of the disease.

Supporting information

S1 Table. Cuproptosis-related genes. The Cuproptosis-related genes used in the manuscript. (DOCX)

S2 Table. The intersection genes of WGCNA and PPI. The intersection genes obtained from WGCNA and PPI. (DOCX)

S3 Table. Enriched KEGG pathways. The 7 pathways significantly enriched in the KEGG analysis. (DOCX)

S4 Table. The significantly different genes in the validation dataset. The significantly differentially expressed genes obtained in the validation dataset. (DOCX)

Acknowledgments

We thank all authors for their contributions to this research.

Additional information

Correspondence and requests for materials should be addressed to J.C. and L.W.

Author Contributions

Conceptualization: Jiajun Chen.

Data curation: Jia Li.

Software: Xiaochen Pei.

Writing – original draft: Jia Song.

Writing – review & editing: Lin Wang.

References

1. Tolosa E, Garrido A, Scholz SW, Poewe W. Challenges in the diagnosis of Parkinson's disease. *Lancet Neurol.* 2021; 20(5): 385–397. [https://doi.org/10.1016/S1474-4422\(21\)00030-2](https://doi.org/10.1016/S1474-4422(21)00030-2) PMID: 33894193
2. Vijjaratnam N, Simuni T, Bandmann O, Morris HR, Foltynie T. Progress towards therapies for disease modification in Parkinson's disease. *Lancet Neurol.* 2021; 20(7): 559–572. [https://doi.org/10.1016/S1474-4422\(21\)00061-2](https://doi.org/10.1016/S1474-4422(21)00061-2) PMID: 34146514
3. Bloem BR, Okun MS, Klein C. Parkinson's disease. *Lancet.* 2021; 397 (10291): 2284–2303. [https://doi.org/10.1016/S0140-6736\(21\)00218-X](https://doi.org/10.1016/S0140-6736(21)00218-X) PMID: 33848468
4. Dinter E, Saridaki T, Diederichs L, Reichmann H, Falkenburger BH. Parkinson's disease and translational research. *Translational Neurodegeneration.* 2020; 9(1): 43. <https://doi.org/10.1186/s40035-020-00223-0> PMID: 33256849
5. Barnham KJ, Bush AI. Metals in Alzheimer's and Parkinson's diseases. *Curr Opin Chem Biol.* 2008; 12(2): 222–228. <https://doi.org/10.1016/j.cbpa.2008.02.019> PMID: 18342639
6. Subramaniam SR, Chesselet M-F. Mitochondrial dysfunction and oxidative stress in Parkinson's disease. *Prog Neurobiol.* 2013;106–107: 17–32. <https://doi.org/10.1016/j.pneurobio.2013.04.004> PMID: 23643800
7. Eldeeb MA, Thomas RA, Ragheb MA, Fallahi A, Fon EA. Mitochondrial quality control in health and in Parkinson's disease. *Physiol Rev.* 2022; 102(4): 1721–1755. <https://doi.org/10.1152/physrev.00041.2021> PMID: 35466694
8. Herrero M-T, Estrada C, Maatouk L, Vyas S. Inflammation in Parkinson's disease: role of glucocorticoids. *Front Neuroanat.* 2015; 9: 32. <https://doi.org/10.3389/fnana.2015.00032> PMID: 25883554
9. Blauwendraat C, Nalls MA, Singleton AB. The genetic architecture of Parkinson's disease. *Lancet Neurol.* 2020; 19(2): 170–178. [https://doi.org/10.1016/S1474-4422\(19\)30287-X](https://doi.org/10.1016/S1474-4422(19)30287-X) PMID: 31521533
10. Gaier ED, Eipper BA, Mains RE. Copper signaling in the mammalian nervous system: synaptic effects. *J Neurosci Res.* 2013; 91(1): 2–19. <https://doi.org/10.1002/jnr.23143> PMID: 23115049
11. Gromadzka G, Tarnacka B, Flaga A, Adamczyk A. Copper Dyshomeostasis in Neurodegenerative Diseases-Therapeutic Implications. *Int J Mol Sci.* 2020; 21(23): E9259. <https://doi.org/10.3390/ijms21239259> PMID: 33291628
12. Giampietro R, Spinelli F, Contino M, Colabufo NA. The Pivotal Role of Copper in Neurodegeneration: A New Strategy for the Therapy of Neurodegenerative Disorders. *Mol Pharm.* 2018; 15(3): 808–820. <https://doi.org/10.1021/acs.molpharmaceut.7b00841> PMID: 29323501
13. Deas E, Cremades N, Angelova PR, Ludtmann MHR, Yao Z, Chen S, et al. Alpha-Synuclein Oligomers Interact with Metal Ions to Induce Oxidative Stress and Neuronal Death in Parkinson's Disease. *Antioxid Redox Signal.* 2016; 24(7): 376–391. <https://doi.org/10.1089/ars.2015.6343> PMID: 26564470
14. Tsvetkov P, Coy S, Petrova B, Dreishpoon M, Verma A, Abdusamad M, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science.* 2022; 375(6586): 1254–1261. <https://doi.org/10.1126/science.abf0529> PMID: 35298263
15. Malpartida AB, Williamson M, Narendra DP, Wade-Martins R, Ryan BJ. Mitochondrial Dysfunction and Mitophagy in Parkinson's Disease: From Mechanism to Therapy. *Trends Biochem Sci.* 2021; 46(4): 329–343. <https://doi.org/10.1016/j.tibs.2020.11.007> PMID: 33323315
16. Okarmus J, Havelund JF, Ryding M, Schmidt SI, Bogetofte H, Heon-Roberts R, et al. Identification of bioactive metabolites in human iPSC-derived dopaminergic neurons with PARK2 mutation: Altered

- mitochondrial and energy metabolism. *Stem cell reports*. 2021; 16(6):1510–1526. <https://doi.org/10.1016/j.stemcr.2021.04.022> PMID: 34048689
17. Lai Y, Lin C, Lin X, Wu L, Zhao Y, Lin F. Identification and immunological characterization of cuproptosis-related molecular clusters in Alzheimer's disease. *Front Aging Neurosci*. 2022; 14: 932676. <https://doi.org/10.3389/fnagi.2022.932676> PMID: 35966780
 18. Wu J, Qin C, Cai Y, Zhou J, Xu D, Lei Y, et al. Machine learning screening for Parkinson's disease-related cuproptosis-related typing development and validation and exploration of personalized drugs for cuproptosis genes. *Ann Transl Med*. 2023; 11(1): 11. <https://doi.org/10.21037/atm-22-5756> PMID: 36760248
 19. Zhang M, Meng W, Liu C, Wang H, Li R, Wang Q, et al. Identification of Cuproptosis Clusters and Integrative Analyses in Parkinson's Disease. *Brain Sci*. 2023; 13(7): 1015. <https://doi.org/10.3390/brainsci13071015> PMID: 37508947
 20. Zhang H, Nagai J, Hao L, Jiang X. Identification of Key Genes and Immunological Features Associated with Copper Metabolism in Parkinson's Disease by Bioinformatics Analysis. *Mol Neurobiol*. Published online September 2, 2023. <https://doi.org/10.1007/s12035-023-03565-8> PMID: 37659036
 21. Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau W-C, Ledoux P, et al. NCBI GEO: mining millions of expression profiles—database and tools. *Nucleic Acids Res*. 2005; 33: D562–566. <https://doi.org/10.1093/nar/gki022> PMID: 15608262
 22. Zhang L, Zhang Y, Bao J, Gao W, Wang D, Pan H. Cuproptosis Combined with lncRNAs Predicts the Prognosis and Immune Microenvironment of Breast Cancer. *Comput Math Methods Med*. 2022; 2022: 5422698. <https://doi.org/10.1155/2022/5422698> PMID: 36213577
 23. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015; 43: D447–452. <https://doi.org/10.1093/nar/gku1003> PMID: 25352553
 24. Smyth GK. limma: Linear Models for Microarray Data. In: Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S, editors. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. New York, NY: Springer; 2005. pp. 397–420. https://doi.org/10.1007/0-387-29362-0_23
 25. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003; 13(11): 2498–2504. <https://doi.org/10.1101/gr.1239303> PMID: 14597658
 26. Tang Y, Li M, Wang J, Pan Y, Wu F-X. CytoNCA: a cytoscape plugin for centrality analysis and evaluation of protein interaction networks. *Biosystems*. 2015; 127: 67–72. <https://doi.org/10.1016/j.biosystems.2014.11.005> PMID: 25451770
 27. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008; 9: 559. <https://doi.org/10.1186/1471-2105-9-559> PMID: 19114008
 28. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000; 25(1): 25–29. <https://doi.org/10.1038/75556> PMID: 10802651
 29. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000; 28(1): 27–30. <https://doi.org/10.1093/nar/28.1.27> PMID: 10592173
 30. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012; 16(5): 284–287. <https://doi.org/10.1089/omi.2011.0118> PMID: 22455463
 31. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011; 12: 77. <https://doi.org/10.1186/1471-2105-12-77> PMID: 21414208
 32. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res*. 2010; 38: W214–220. <https://doi.org/10.1093/nar/gkq537> PMID: 20576703
 33. Huang Z, Shi J, Gao Y, Cui C, Zhang S, Li J, et al. HMDD v3.0: a database for experimentally supported human microRNA-disease associations. *Nucleic Acids Res*. 2019; 47: D1013–D1017. <https://doi.org/10.1093/nar/gky1010> PMID: 30364956
 34. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: An online resource for prediction of microRNA binding sites. *PLoS One*. 2018; 13(10): e0206239. <https://doi.org/10.1371/journal.pone.0206239> PMID: 30335862
 35. Vlachos IS, Zagganas K, Paraskevopoulou MD, Georgakilas G, Karagkouni D, Vergoulis T, et al. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. *Nucleic Acids Res*. 2015; 43: W460–466. <https://doi.org/10.1093/nar/gkv403> PMID: 25977294

36. Montes S, Rivera-Mancia S, Diaz-Ruiz A, Tristan-Lopez L, Rios C. Copper and copper proteins in Parkinson's disease. *Oxid Med Cell Longev*. 2014; 2014: 147251. <https://doi.org/10.1155/2014/147251> PMID: 24672633
37. Bisaglia M, Bubacco L. Copper Ions and Parkinson's Disease: Why Is Homeostasis So Relevant? *Biomolecules*. 2020; 10(2): 195. <https://doi.org/10.3390/biom10020195> PMID: 32013126
38. Hung LW, Villemagne VL, Cheng L, Sherratt NA, Ayton S, White AR, et al. The hypoxia imaging agent Cull(atm) is neuroprotective and improves motor and cognitive functions in multiple animal models of Parkinson's disease. *J Exp Med*. 2012; 209(4): 837–854. <https://doi.org/10.1084/jem.20112285> PMID: 22473957
39. Genoud S, Senior AM, Hare DJ, Double KL. Meta-Analysis of Copper and Iron in Parkinson's Disease Brain and Biofluids. *Mov Disord*. 2020; 35(4): 662–671. <https://doi.org/10.1002/mds.27947> PMID: 31889341
40. Tejera PO'Mahony DS, Owen CA, Wei Y, Wang Z, Gupta K, et al. Functional characterization of polymorphisms in the peptidase inhibitor 3 (elafin) gene and validation of their contribution to risk of acute respiratory distress syndrome. *Am J Respir Cell Mol Biol*. 2014; 51(2): 262–272. <https://doi.org/10.1165/rcmb.2013-0238OC> PMID: 24617927
41. Caruso JA, Karakas C, Zhang J, Yi M, Albarracin C, Sahin A, et al. Elafin is downregulated during breast and ovarian tumorigenesis but its residual expression predicts recurrence. *Breast Cancer Res*. 2014; 16(6): 3417. <https://doi.org/10.1186/s13058-014-0497-4> PMID: 25551582
42. Bartl M, Dakna M, Schade S, Otte B, Wicke T, Lang E, et al. Blood Markers of Inflammation, Neurodegeneration, and Cardiovascular Risk in Early Parkinson's Disease. *Mov Disord*. 2023; 38(1): 68–81. <https://doi.org/10.1002/mds.29257> PMID: 36267007
43. Schipanski A, Lange S, Segref A, Gutschmidt A, Lomas DA, Miranda E, et al. A novel interaction between aging and ER overload in a protein conformational dementia. *Genetics*. 2013; 193(3): 865–876. <https://doi.org/10.1534/genetics.112.149088> PMID: 23335331
44. Roussel BD, Kruppa AJ, Miranda E, Crowther DC, Lomas DA, Marciniak SJ. Endoplasmic reticulum dysfunction in neurological disease. *Lancet Neurol*. 2013; 12(1): 105–118. [https://doi.org/10.1016/S1474-4422\(12\)70238-7](https://doi.org/10.1016/S1474-4422(12)70238-7) PMID: 23237905
45. Zattoni M, Mearelli M, Vanni S, Colini Baldeschi A, Tran TH, Ferracin C, et al. Serpin Signatures in Prion and Alzheimer's Diseases. *Mol Neurobiol*. 2022; 59(6): 3778–3799. <https://doi.org/10.1007/s12035-022-02817-3> PMID: 35416570
46. Kinghorn KJ, Crowther DC, Sharp LK, Nerelius C, Davis RL, Chang HT, et al. Neuroserpin binds Aβeta and is a neuroprotective component of amyloid plaques in Alzheimer disease. *J Biol Chem*. 2006; 281(39): 29268–29277. <https://doi.org/10.1074/jbc.M600690200> PMID: 16849336
47. Fabbro S, Schaller K, Seeds NW. Amyloid-beta levels are significantly reduced and spatial memory defects are rescued in a novel neuroserpin-deficient Alzheimer's disease transgenic mouse model. *J Neurochem*. 2011; 118(5): 928–938. <https://doi.org/10.1111/j.1471-4159.2011.07359.x> PMID: 21689108
48. Stojkowska I, Wani WY, Zunke F, Belur NR, Pavlenko EA, Mwenda N, et al. Rescue of α-synuclein aggregation in Parkinson's patient neurons by synergistic enhancement of ER proteostasis and protein trafficking. *Neuron*. 2022; 110(3): 436–451.e11. <https://doi.org/10.1016/j.neuron.2021.10.032> PMID: 34793693
49. Joshi R, Heinz A, Fan Q, Guo S, Monia B, Schmelzer CEH, et al. Role for Cela1 in Postnatal Lung Remodeling and Alpha-1 Antitrypsin-Deficient Emphysema. *Am J Respir Cell Mol Biol*. 2018; 59(2): 167–178. <https://doi.org/10.1165/rcmb.2017-0361OC> PMID: 29420065
50. Wang W, Zhang L, Sun J, Zhao Q, Shuai J. Predicting the potential human lncRNA-miRNA interactions based on graph convolution network with conditional random field. *Brief Bioinform*. 2022; 23(6): bbac463. <https://doi.org/10.1093/bib/bbac463> PMID: 36305458
51. Zhang L, Yang P, Feng H, Zhao Q, Liu H. Using Network Distance Analysis to Predict lncRNA-miRNA Interactions. *Interdiscip Sci*. 2021; 13(3): 535–545. <https://doi.org/10.1007/s12539-021-00458-z> PMID: 34232474
52. Hu H, Feng Z, Lin H, Cheng J, Lyu J, Zhang Y, et al. Gene function and cell surface protein association analysis based on single-cell multiomics data. *Comput Biol Med*. 2023; 157: 106733. <https://doi.org/10.1016/j.combiomed.2023.106733> PMID: 36924730
53. Gao H, Sun J, Wang Y, Lu Y, Liu L, Zhao Q, et al. Predicting metabolite-disease associations based on auto-encoder and non-negative matrix factorization. *Brief Bioinform*. 2023; 24(5): bbad259. <https://doi.org/10.1093/bib/bbad259> PMID: 37466194
54. Sun F, Sun J, Zhao Q. A deep learning method for predicting metabolite-disease associations via graph neural network. *Brief Bioinform*. 2022; 23(4): bbac266. <https://doi.org/10.1093/bib/bbac266> PMID: 35817399

55. Chen Z, Zhang L, Sun J, Meng R, Yin S, Zhao Q. DCAMCP: A deep learning model based on capsule network and attention mechanism for molecular carcinogenicity prediction. *J Cell Mol Med.* 2023; 27(20): 3117–3126. <https://doi.org/10.1111/jcmm.17889> PMID: 37525507
56. Wang T, Sun J, Zhao Q. Investigating cardiotoxicity related with hERG channel blockers using molecular fingerprints and graph attention mechanism. *Comput Biol Med.* 2023; 153: 106464. <https://doi.org/10.1016/j.combiomed.2022.106464> PMID: 36584603
57. Meng R, Yin S, Sun J, Hu H, Zhao Q. sCAAGA: Single cell data analysis framework using asymmetric autoencoder with gene attention. *Comput Biol Med.* 2023; 165: 107414. <https://doi.org/10.1016/j.combiomed.2023.107414> PMID: 37660567
58. Xie S, Niu W, Xu F, Wang Y, Hu S, Niu C. Differential expression and significance of miRNAs in plasma extracellular vesicles of patients with Parkinson's disease. *Int J Neurosci.* 2022; 132(7): 673–688. <https://doi.org/10.1080/00207454.2020.1835899> PMID: 33045885
59. Zhang Y, Zhou Q, Lu L, Su Y, Shi W, Zhang H, et al. Copper Induces Cognitive Impairment in Mice via Modulation of Cuproptosis and CREB Signaling. *Nutrients.* 2023; 15(4): 972. <https://doi.org/10.3390/nu15040972> PMID: 36839332
60. Li X, Zhong C-Q, Wu R, Xu X, Yang Z-H, Cai S, et al. RIP1-dependent linear and nonlinear recruitments of caspase-8 and RIP3 respectively to necrosome specify distinct cell death outcomes. *Protein Cell.* 2021; 12(11): 858–876. <https://doi.org/10.1007/s13238-020-00810-x> PMID: 33389663
61. Xu F, Miao D, Li W, Jin J, Liu Z, Shen C, et al. Specificity and competition of mRNAs dominate droplet pattern in protein phase separation. 2023; 5(2): 023159. <https://doi.org/10.1103/PhysRevResearch.5.023159>
62. Jin J, Xu F, Liu Z, Shuai J, Li X. Quantifying the underlying landscape, entropy production and biological path of the cell fate decision between apoptosis and pyroptosis. *Chaos, Solitons & Fractals.* 2024; 178: 114328. <https://doi.org/10.1016/j.chaos.2023.114328>