RESEARCH



An integrated bioinformatics and machine learning approach to identifying biomarkers connecting parkinson's disease with purine metabolism-related genes

Yao Wang¹⁺, Dongchuan Wu²⁺, Man Zheng³ and Tiantian Yang^{4*}

Abstract

Background Parkinson's disease (PD), a prevalent neurodegenerative disorder in the aging population, poses significant challenges in unraveling its pathogenesis and progression. A key area of investigation is the disruption of oncological metabolic networks in PD, where diseased cells display distinct metabolic profiles compared to healthy counterparts. Of particular interest are Purine Metabolism Genes (PMGs), which play a pivotal role in nucleic acid synthesis.

Methods In this study, bioinformatics analyses were employed to identify and validate PMGs associated with PD. A set of 20 candidate PMGs underwent differential expression analysis. GSEA and GSVA were conducted to explore the biological roles and pathways of these PMGs. Lasso regression and SVM-RFE methods were applied to identify hub genes and assess the diagnostic efficacy of the nine PMGs in distinguishing PD. The correlation between these hub PMGs and clinical characteristics was also explored. Validation of the expression levels of the nine identified PMGs was performed using the GSE6613 and GSE7621 datasets.

Results The study identified nine PMGs related to PD: NME7, PKM, RRM2, POLR3 C, POLA1, PDE6 C, PDE9 A, PDE11 A, and AMPD1. Biological function analysis highlighted their involvement in processes like neutrophil activation and immune response. The diagnostic potential of these nine PMGs in differentiating PD was found to be substantial.

Conclusions This investigation successfully identified nine PMGs associated with PD, providing valuable insights into potential novel biomarkers for this condition. These findings contribute to a deeper understanding of PD's pathogenesis and may aid in monitoring its progression, offering a new perspective in the study of neurodegenerative diseases.

Keywords Parkinson's disease (PD), Purine Metabolism Genes (PMGs), Biomarker candidates, POLA1, POLR3 C

 $^{\dagger}\mathrm{Yao}$ Wang and Dongchuan Wu contributed equally this work as the first author.

*Correspondence: Tiantian Yang ytt@bucm.edu.cn Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Parkinson's disease (PD), the second-most common neurodegenerative condition after Alzheimer's disease, affects approximately 1.2% of individuals over the age of 65. This disorder predominantly targets the older population, typically manifesting around the age of 60, with aging as the primary risk factor [1]. Characterized by a profound impairment in motor coordination, PD arises from the degeneration of dopaminergic neurons in the substantia nigra (SN) [2]. Clinically, it is marked by distinct symptoms including resting tremors, bradykinesia, rigidity, and postural-gait abnormalities [3]. PD also encompasses a spectrum of additional motor dysfunctions such as altered gait and posture, difficulties in speech and swallowing, and expressive changes [4]. Recent advancements have led to the recognition of nonmotor symptoms in PD. These non-motor symptoms often have a more pronounced impact on patient guality of life than the motor symptoms, underscoring the critical need for research focusing on their prevention [5]. The etiology of PD is multifaceted, with genetic predisposition, environmental factors, aging, and oxidative stress all contributing to the degeneration of dopaminergic neurons [6]. This review aims to provide a comprehensive understanding of PD, with a particular focus on the unresolved complexities of PD-associated inflammation [7]. There is a pressing need for continued research to elucidate the intricate role of immune-inflammatory mechanisms in PD pathophysiology. Insights from such investigations are essential for the development of accurate diagnostic biomarkers and targeted therapeutic strategies [8]. In this context, the early identification of PD-specific molecular biomarkers is critical [9]. Such biomarkers could enable timely intervention, ideally before the onset of motor symptoms, thereby transforming the clinical management and prognosis of PD. This strategy holds the potential to shift the therapeutic paradigm toward early detection and intervention, ultimately improving disease outcomes.

The metabolic reprogramming's results in a distinct metabolic phenotype within tumor cells, reshaping the immune microenvironment. The immune microenvironment, characterized by a heterogeneous mix of cell types and challenged by poor oxygen and nutrient supply, plays a crucial role in cancer progression [10]. Recent advancements highlight the significance of nontumoral immune infiltration in the TME, with growing evidence linking immune response to profound changes in tissue metabolism [11]. These include nutrient depletion, increased oxygen consumption, and the production of reactive species, which collectively influence immune cell functionality and maturation. This understanding opens avenues for metabolic interventions to augment the efficacy of immunotherapies [12]. Purines, essential components of DNA, RNA, and key biomolecules like ATP and NADH, are central to various cellular functions, including energy production, signal transduction, and fatty acid biosynthesis [13]. Moreover, purines significantly influence immune responses and host-pathogen interactions [14]. Mammalian cells predominantly satisfy their purine requirements through the salvage pathway, but in rapidly proliferating cells, such as tumor cells, an enhanced need for purines is met by upregulating de novo synthesis [15]. Historically, purine antimetabolites were among the first anticancer agents and continue to play a vital role in treating various leukemias and non-neoplastic diseases by inhibiting DNA synthesis and cell growth [16]. The discovery of purinosomes, intricately linked to the cell cycle, offers a novel therapeutic target in purine metabolism [17]. While the integration of purine metabolic strategies with immunotherapy shows promise, especially in PD management, the role of purine metabolism in immunogenicity and immunotherapy remains largely unexplored. This study endeavors to bridge this knowledge gap, aiming for an in-depth evaluation of PMGs in the context of immunotherapy for PD. This approach could revolutionize the understanding and treatment of PD, paving the way for innovative therapeutic strategies.

The PD Initiative, by integrating comprehensive transcriptome sequencing data with detailed clinical annotations, offers a unique and invaluable resource for exploring the transcriptional intricacies and molecular pathways underlying PD. Bioinformatics analyses of these extensive datasets have provided critical insights into the complex pathophysiology of PD [18–20]. However, a notable gap in the current research lies in the application of bioinformatics to investigate the role of PMGs in PD. To address this gap, our study leverages PD-related GEO datasets to investigate the significance and impact of PMGs in PD pathogenesis, as illustrated in Fig. 1.

Materials and methods

Raw Data and Differentially Expressed Genes (DEGs)

We utilized two foundational datasets from the GEO series, GSE6613 and GSE7621. GSE6613 was used for training, while GSE7621 was reserved for validation. Additionally, the MSigDB provided a comprehensive list of PMGs (Table S1). mRNA profiles were extracted using Perl scripts to match and sort transcriptional data. Following normalization, DEGs among the PMGs were identified using criteria: FDR <0.05 and $|log2 \text{ FC}| \ge 1$. Pearson's correlation coefficient was then employed, using the corrplot package in R.



Fig. 1 Framework. To advance our understanding of PD, we conducted a comprehensive analysis using patient-derived datasets from the GEO repository. Our primary cohort included the GSE6613 and GSE7621 datasets, with the GSE7621 dataset employed for validation. By applying a rigorous matching strategy for PMGs, we performed differential expression analyses and constructed a prognostic risk model. This approach identified a distinct subset of PMGs with prognostic significance in PD, highlighting their potential as candidate biomarkers. To further explore the functional roles of these genes, we conducted an extensive array of bioinformatics analyses, including GO, KEGG, and GSEA. These analyses were supplemented with data from multiple databases, offering a multidimensional view of the implicated PMGs and their involvement in cellular processes, signaling pathways, and gene regulatory networks

Building a model, immune cell infiltration and functional enrichment analysis

To elucidate the biological significance and pathway involvement of differentially expressed genes (DEGs), we conducted Gene Ontology (GO) and KEGG pathway analyses. Using the R platform, we assessed the impact of differentially expressed PMGs on biological processes (BP), molecular functions (MF), and cellular components (CC), providing a detailed functional characterization of these genes. Model refinement was accomplished through Lasso regression using the glmnet package [21] with repeated K-fold cross-validation, enabling the selection of an optimal penalty parameter to maximize model accuracy and predictive strength. To further validate the model, we implemented the Support Vector Machine-Recursive Feature Elimination (SVM-RFE) algorithm via the e1071 package of R.4.3 [22], constructing a machine learning model with high precision. Cross-validation was instrumental in evaluating model performance, minimizing error rates, and enhancing predictive accuracy, thereby ensuring the model's robustness. Key gene analysis and disease classification were conducted using the ggplot2 package [23], allowing us to identify and visualize critical genes implicated in disease pathogenesis. Additionally, immune cell composition was analyzed using the CIBERSORT algorithm [24], offering comprehensive insights into the immune landscape associated with the disease and highlighting potential immune regulatory mechanisms.

Gene set enrichment and variation analyses and drug-gene interaction insights

To investigate functional dynamics and pathway alterations across diverse samples, we conducted Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA). Leveraging the R platform, we evaluated the influence of differentially expressed PMGs on BP, MF, and CC, and signaling pathways, providing detailed insights into their roles in disease mechanisms. Given the pivotal role of validated biomarkers in shaping therapeutic strategies, accurate drug prediction remains essential.

Construction of an mRNA-miRNA-IncRNA Network

Non-coding RNA transcripts play pivotal roles in shaping the genetic regulatory network. miRNAs modulate gene expression by promoting mRNA degradation and inhibiting translation, while lncRNAs, typically over 200 nucleotides in length, influence diverse cellular processes through chromatin modification, transcriptional regulation, and interference mechanisms. Recent studies underscore the intricate crosstalk between miRNAs and lncRNAs, leading to competitive binding interactions with other regulatory molecules. This phenomenon, termed the ceRNA network, reveals how lncRNAs regulate gene expression by sequestering miRNAs, thereby modulating their activity. To investigate these interactions, we curated target gene information from miRTar-Base and PrognoScan, databases that provide validated miRNA-lncRNA-target relationships.

Mendelian randomization analysis

To ensure the independence of exposure and outcome variables in our genome-wide association study (GWAS) summary data, we conducted an association analysis using the TwoSampleMR package in R. We designated PMG-related expression as the exposure and PD as the outcome to investigate potential causal relationships. The analysis comprised three key steps: (1)Instrumental Variable (IV) Selection: PMG-related expressions were filtered using a significance threshold of $P < 5 \times 10^{-8}$ to identify strongly associated SNPs. (2) Independence Configuration: Linkage disequilibrium (LD) between SNPs was calculated using the PLINK clustering method, excluding SNPs with an LD coefficient ($r^2 > 0.001$) within a 10,000 kb window to ensure SNP independence and

minimize pleiotropic bias. (3) Statistical Strength Assessment: The robustness of instrumental variables was evaluated using the F-statistic (F = β^2 /SE²), with F < 10 indicating insufficient strength to mitigate confounding. Following SNP identification, the harmonise_data function within TwoSampleMR was employed to align allelic directions between exposure and outcome, excluding incompatible SNPs. Causal inference was performed using the inverse variance-weighted (IVW) method, which leverages the variance of instrumental variables as weights to estimate causal effects, thereby providing insights into the genetic architecture underlying PD susceptibility.

Results

Identification of degs and enrichment analysis of PMGs

Among the 20 examined PMGs, several exhibited significant differences in expression levels. Furthermore, gene clustering analysis revealed distinct clusters in the PD and control groups. Notable PMGs in the PD group included AMPD1, PDE1 C, XDH, ADCY10, ENPP1, RRM2, ADCY2, while control group PMGs comprised NME7, PKM, IMPDH2, POLR2L, ADPRM, ENTPD4 (Fig. 2a). Correlation analysis was conducted among these PMGs (Fig. 2b) (Table S2). The MF category primarily involved phosphoric magnesium ion binding (GO:0000287), lyase activity (GO:0016829), nucleotidyltransferase activity (GO:0016779). The CC category was mainly associated with neuronal cell body (GO:0043025), vesicle lumen (GO:0031983), cytoplasmic vesicle lumen (GO:0060205). The BP category included neutrophil activation (GO:0042119), neutrophil mediated immunity (GO:0002446), neutrophil activation involved in immune response (GO:0002283). Pathway enrichment analysis revealed that DEGs in PD are predominantly involved in purine metabolism (hsa00230), pyrimidine metabolism (hsa00240), and drug metabolism – other enzymes (hsa00983), highlighting their potential roles in PD pathogenesis. Purine metabolism is central to ATP and GTP production, essential for maintaining neuronal energy homeostasis and neurotransmission. Disruption of this pathway may impair mitochondrial function, exacerbate oxidative stress, and contribute to dopaminergic neuron degeneration in PD. Pyrimidine metabolism is crucial for nucleotide biosynthesis and RNA processing. Dysregulation of this pathway could compromise neuronal repair mechanisms and synaptic plasticity, thereby accelerating neurodegeneration. Alterations in drug metabolism pathways may influence the pharmacokinetics and efficacy of PD therapeutics. Variations in drug-metabolizing enzyme activity could lead to differential treatment responses and reduced therapeutic effectiveness. These findings suggest that metabolic reprogramming



Fig. 2 DEGs and Enrichment Analysis. **a** Analysis of difference. **b** Analysis of correlation. **c**: GO. (d): KEGG. Bubble graph for GO enrichment (the bigger bubble means the more genes enriched, and the increasing depth of red means the differences were more obvious; q-value: the adjusted p-value); The GO circle shows the scatter map of the logFC of the specified gene. Barplot graph for KEGG pathways (the longer bar means the more genes enriched, and the increasing depth of red means the differences were more obvious); The KEGG circle shows the scatter map of the logFC of the specified gene. Barplot graph for KEGG circle shows the scatter map of the logFC of the specified gene. The higher the Z-score value indicated, the higher expression of the enriched pathway

in PD involves coordinated dysregulation of nucleotide metabolism and drug processing, which may collectively exacerbate disease progression and treatment resistance. Further investigation into these pathways could uncover novel therapeutic targets and improve patient outcomes (Fig. 2c-d and Table S3a-b).

Model construction

We constructed a gene signature through LASSO and Cox regression analysis, selecting the optimal penalty value to maximize predictive performance (Fig. 3a-b). To validate the model's precision and reliability, we developed a machine learning framework using SVM-RFE, which achieved a notable accuracy of 0.769 with a low error rate of 0.231 (Fig. 3c-d). Moreover, the overlap of the nine PMGs identified through both LASSO and SVM-RFE demonstrated significant concordance, reinforcing the robustness of the selected signature (Fig. 3e). Performance evaluation of the model for the nine hub genes revealed consistently high predictive accuracy, as indicated by the area under the curve (AUC) values: NME7 (0.664), PKM (0.640), RRM2 (0.647), POLR3 C (0.662), POLA1 (0.663), PDE6 C (0.630), PDE9 A (0.663), PDE11 A (0.639), and AMPD1 (0.632) (Fig. 3f). Notably, the predictive strength of the model was further corroborated by an AUC of 0.912 (95% confidence interval (CI): 0.840-0.968) in the independent validation dataset GSE7621, underscoring the model's high accuracy and generalizability (Fig. 3g) (Table 1 and S4). Regarding the overall model performance, the AUC value of 0.769

Table 1 The characteristics of	of model
--------------------------------	----------

Label	LASSO	SVM-RFE
Sensitivity	0.416667	0.500000
Specificity	0.800000	1.000000
Pos Pred Value	0.555556	1.000000
Neg Pred Value	0.695652	0.769231
Precision	0.555556	1.000000
Recall	0.416667	0.500000
F1	0.476190	0.666667
Prevalence	0.375000	0.375000
Detection Rate	0.156250	0.187500
Detection Prevalence	0.281250	0.187500
Balanced Accuracy	0.608333	0.750000

reflects the model's robust predictive capacity. While this value may be perceived as modest, it is essential to account for the inherent variability introduced by individual genetic differences, which may partially constrain the maximum achievable accuracy.

GSEA and **GSVA**

In terms of GO analysis, POLA1 was found to be associated with CC photoreceptor outer segment membrane, BP regulation of transposition, BP steroid catabolic process. On the other hand, POLR3 C was primarily involved in the BP cellular response to antibiotic, CC fascia adherens, BP regulation of skeletal muscle cell differentiation



Fig. 3 The development of the PMGs signature. a Regression of LASSO. b Cross-validation. c-d Accuracy and error. e Venn. f-g AUC of hub gene and train group. Regarding the overall model performance, the AUC value of 0.769 reflects the model's robust predictive capacity

(Fig. 4a). In KEGG analysis, POLA1 was mainly associated with KEGG cytokine cytokine receptor interaction, KEGG gnrh signaling pathway, KEGG proximal tubule bicarbonate reclamation, while POLR3 C was involved in KEGG renin angiotensin system, KEGG base excision repair, KEGG glycosaminoglycan biosynthesis heparan sulfate (Fig. 4b) (Table S5). In the GO analysis, POLA1 was primarily associated with BP regulation of response to drug, BP negative regulation of cell chemotaxis to fibroblast growth factor, CC egg coat, BP transposition rna mediated, CC fancm mhf complex, BP threonine catabolic process, BP meiotic dna double strand break formation, MF glycine n acyltransferase activity. POLR3 C was mainly involved in the BP threonine catabolic process, BP rna 3 uridylation, MF protein disaggregase activity, CC fhf complex, BP sperm mitochondrial sheath assembly, MF lysophosphatidic acid phosphatase activity, MF udp xylosyltransferase activity, CC iga immunoglobulin complex, BP phosphagen metabolic process, BP negative regulation of myoblast proliferation (Fig. 4c). In terms of KEGG analysis, POLA1 was mainly associated with asthma, steroid biosynthesis, non homologous end joining, nitrogen metabolism, ether lipid metabolism, alpha linolenic acid metabolism, linoleic acid metabolism, proximal tubule bicarbonate reclamation. POLR3 C was involved in glycosaminoglycan biosynthesis heparan sulfate, glycosaminoglycan biosynthesis keratan sulfate, asthma, steroid biosynthesis, dorso ventral axis formation, folate biosynthesis (Fig. 4d).

Analysis of immune cells

This section of the study delves into the intricate role of the immune microenvironment in the onset and progression of PD. Recognizing the significance of immune cell dynamics, a detailed analysis was conducted to elucidate their expression patterns and interactions in PD. To visually represent these patterns, a vioplot was created, which effectively highlighted the differential expression of various immune cells between control and PD groups. Notably, a vioplot was created to display the expression patterns of T cells CD4 memory activated, T cells follicular helper, T cells gamma delta, Dendritic cells resting, which were highly expressed in the control group. While, T cells regulatory (Tregs), Macrophages M0, Mast cells resting, Mast cells activated, Neutrophils. This elevated expression in the PD group provides critical insights into the immune response mechanisms that might be activated or altered in the context of PD. Further adding to this immunological landscape, a comprehensive correlation analysis was conducted. This analysis aimed to unravel the intricate relationships between the identified genes and various immune cells. Such insights are pivotal for understanding how genetic factors might influence or be influenced by the immune microenvironment in PD. The findings from these analyses, as illustrated in Fig. 5ab, shed light on the complex interplay between immune cells and genetic factors in PD.

miRNA-IncRNA shared genes network

A total of 143 miRNAs and 190 lncRNAs associated with PD were identified from three databases (Table S6a-b). The network consisted of 157 lncRNAs, 131 miRNAs, and some common genes, including the 6 hub genes (PDE11 A, POLA1, PDE9 A, RRM2, NME7, PDE6 C) (Fig. 6).

Model and hub genes external validation

To verify the reliability of our results. We supplemented the expression profile of these genes in other machine learning methods. These methods include. RF, SVM, XGB, and GLM. RF is an ensemble learning method that builds multiple decision trees and combines their outputs to enhance predictive accuracy and control overfitting. By aggregating the results of numerous decision trees, RF improves model stability and reduces variance. SVM is a supervised learning algorithm that identifies the hyperplane that best separates different classes in a high-dimensional space. It is particularly effective in handling complex, non-linear relationships through the use of kernel functions. XGB is a powerful gradient-boosting algorithm that refines prediction accuracy by sequentially correcting the errors of preceding models. It incorporates regularization to prevent overfitting and improve generalization. GLM extends linear regression by allowing for different response distributions and link functions, making it suitable for modeling complex relationships in non-normally distributed data. The boxplots illustrated the residual expression patterns of the identified genes in PD (Fig. 7a-b). Notable differences were observed in the proportions of the four predictive models (Fig. 7c). The diagnostic performance of PMGs in distinguishing PD from control samples demonstrated a high predictive value, with the AUC values as follows: RF: 0.967; SVM: 0.919; XGB: 0.943; GLM: 0.933 (Fig. 7d). To enhance the confidence and prediction accuracy of the model, GSE7621 dataset was used for validation. The GSE7621 analysis further confirming their potential relevance to PD (Fig. 7f).

Mendelian randomization analysis

In examining the direct linkage between the PMGs (POLA1 and POLR3 C) and PD incidence, a forest plot was utilized for visual illustration, revealing a general symmetry in the data. Through sensitivity analysis employing the "leave-one-out" technique, it was determined that the omission of any individual SNP had



Fig. 4 Expression of Immune cells. a Expression of immune cells in different clusters. b Correlation between PMGs and immune cells



Fig. 5 miRNAs-LncRNAsNetwork



Fig. 6 Model verification. **a-b** Residual expression patterns. **c** Model expression patterns (**d**) AUC of model. **e** AUC of test group. **f** Nine hub genes were validated

a minimal effect on the results of the inverse varianceweighted (IVW) analysis, indicating that the remaining SNPs closely mirrored the overall dataset's findings. To further authenticate our outcomes, MR-Egger regression analysis was conducted, bolstering the integrity and reliability of our results and the chosen analytical framework (Fig. 8a-b).



Fig. 7 Model verification. (a-b) Residual expression patterns. c Model expression patterns (d) AUC of model. e AUC of test group. f Nine hub genes were validated

Discussions

PD, a common neurodegenerative disorder, manifests classic motor symptoms such as tremor, akinesia, and bradykinesia, alongside non-motor symptoms like constipation, sleep disturbances, and cognitive deficits [25]. Pathologically characterized by the accumulation of α -synuclein aggregates forming Lewy bodies and the degeneration of dopaminergic neurons in the substantia nigra pars compacta (SNpc) [26], PD's incidence increases with age [27]. However, its neurodegenerative

pathogenesis is only partially understood, attributed to factors including genetic predispositions, oxidative stress, immune dysregulation, mitochondrial dysfunction, and lipid homeostasis disruption [28]. Recent research has focused on the modulation of programmed cell death pathways in tumor biology as a promising strategy for PD therapeutics [29]. This shift has brought metabolic markers, particularly those related to cysteine and nucleotide metabolism, and oncometabolites like 2-hydroxyglutarate, into the spotlight for both diagnostic and therapeutic



Fig. 8 Mendelian Randomization Analysis. a POLA1. b POLR3 C

purposes. Consequently, early diagnosis and precise PD identification are paramount [30]. In this context, the regulation of gene expression is crucial. Purine, a key metabolic component involved in various cell signaling processes, and its metabolism regulators, play a significant role in tumor cell proliferation and treatment resistance [31]. Disruptions in purine nucleotide metabolism, affecting gene and protein expression, can enhance cellular malignancy, invasiveness, and metastasis [32]. While recent research has identified several risk markers in various diseases, their practical application is limited by the absence of comprehensive reviews and large-scale replication. Most studies to date have concentrated on the effects of single purine metabolism regulators in cancer [33]. The collective impact of multiple purine metabolism-related genes in other diseases, including PD, remains underexplored [34]. As our understanding of tumor biology deepens, research is increasingly focusing on non-tumoral aspects. Investigating the diverse purine metabolic patterns in PD could illuminate the role of purine metabolism in PD's progression, offering new potential targets for therapeutic intervention.

In our study, we identified a set of 20 DEGs intricately linked with PMGs in PD. Utilizing a robust methodology that integrates DEG analysis, Lasso regression, and SVM-RFE, we pinpointed nine key PMGs: NME7, PKM, RRM2, POLR3 C, POLA1, PDE6 C, PDE9 A, PDE11 A, and AMPD1. These hub genes demonstrated significant diagnostic potential for PD, a finding supported by external dataset validation, highlighting their pivotal role in the disease's pathogenesis. The identification of these PMGs not only deepens our understanding of PD pathogenesis but also holds tangible clinical relevance. Purine metabolism plays a crucial role in neuroinflammation, mitochondrial function, and oxidative stress, all of which are central to PD's neurodegenerative process. Elevated expression of RRM2 and PKM may reflect heightened nucleotide biosynthesis and altered glycolytic activity, consistent with increased cellular stress and metabolic dysregulation observed in PD [35]. Conversely, dysregulation of phosphodiesterase genes (PDE6 C, PDE9 A, and PDE11 A) implicates impaired cyclic nucleotide signaling, potentially disrupting synaptic plasticity and neurotransmitter homeostasis-pathological hallmarks of PD [36]. Notably, the downregulation of AMPD1 suggests compromised purine salvage, which could exacerbate cellular energy deficits and contribute to neuronal vulnerability. From a diagnostic perspective, the nine PMGs demonstrate substantial potential as biomarkers for early PD detection [10]. Their inclusion in a composite predictive model could enhance diagnostic accuracy, particularly in distinguishing PD from other

neurodegenerative conditions with overlapping clinical features. Moreover, these findings provide a foundation for targeted therapeutic strategies, such as modulating purine metabolism and cyclic nucleotide signaling, which may mitigate disease progression. Further investigation into the mechanistic underpinnings of these PMGs could facilitate the development of precision medicine approaches, ultimately improving patient outcomes in PD. However, our findings also underscore a notable gap in understanding the interactions between these genes and specific transcription factors within the PMGs framework.

A comprehensive literature review identified POLA1 and POLR3 C as pivotal regulators in the complex interplay between PD and PMGs. Further investigation into their biological roles revealed their involvement in metabolic pathways linked to neutrophil activation and immune response. This suggests that PMGs may exert broad regulatory effects on diverse biological processes, particularly those influencing immune-related pathways. Such regulatory impact could significantly shape the pathophysiological progression of PD, positioning these genes as potential targets for therapeutic intervention. Purine metabolism, a cornerstone in maintaining cellular energy homeostasis and supporting cell proliferation, has been increasingly recognized for its critical role in both oncogenesis and metabolic disorders. Central to this metabolic axis is POLA1, which encodes the catalytic subunit of DNA polymerase α , a key enzyme responsible for initiating DNA replication in concert with the Primase complex [37]. While POLA1 has traditionally been considered essential for cellular viability, recent evidence indicates that its partial deficiency is linked to at least two distinct disorders [38]. The first identified syndrome, X-linked reticulate pigmentary disorder (XLPDR, MIM #301,220), is characterized by distinctive skin hyperpigmentation, systemic sterile inflammation, recurrent infections, and unique craniofacial features [39]. Moreover, the integrity of epigenetic regulation during DNA replication-specifically the recycling of parental histones and the deposition of new histones-is tightly orchestrated by a complex network of histone chaperones, remodelers, and binding proteins [40]. Disruption of this finely tuned process can precipitate genomic instability and altered gene expression patterns, underscoring its relevance in both developmental and pathological contexts. In oncological research, Li et al. [41]. identified POLR3 C and KPNA2 as neoantigens associated with poor prognosis in liver hepatocellular carcinoma (LIHC). These neoantigens are linked to heightened infiltration of antigen-presenting cells, implicating them in tumor immune evasion and progression. This highlights the potential dual role of POLR3 C in immune modulation

and oncogenesis, reinforcing the need for further investigation into its mechanistic underpinnings and therapeutic potential.

Recent advances in molecular genetics have uncovered a complex and tissue-specific involvement of RNA polymerase (Pol) III, traditionally known for its role in transcribing small untranslated RNAs vital for RNA maturation and translation, in various inherited diseases [42]. This paradigm shift from an anticipated generalized cellular dysfunction reveals a more intricate involvement of Pol III in human diseases [41]. Notably, mutations in the POLR3 A, POLR3 C, POLR3E, and POLR3 F subunits are now linked with heightened susceptibility to varicella zoster virus-induced encephalitis and pneumonitis [43]. Furthermore, an expanding array of mutations in POLR3 A, POLR3B, POLR1 C, and POLR3 K subunits is associated with a range of neurodegenerative diseases, including notably hypomyelinating leukodystrophy. Additional rare disorders have been traced back to mutations in POLR3H, POLR3GL, and the BRF1 component of the TFIIIB transcription initiation factor [44]. Although the correlation between these genetic variations and the manifestation of diseases is clear, the exact molecular mechanisms underlying their pathogenesis remain a subject of ongoing investigation. Our analysis of the GSE7621 dataset suggests that purine-related features could serve as potential prognostic markers in PD, indicating a nascent yet promising area of genomic research. The identification of POLA1 and POLR3 C as potential molecular mediators of PD pathogenesis underscores their diagnostic and therapeutic relevance. Elevated or suppressed expression of these genes may serve as early biomarkers for PD, aiding in differential diagnosis and risk stratification. Furthermore, targeting the DNA replication and RNA transcription pathways modulated by these genes could open novel therapeutic avenues. For instance, pharmacological modulation of POLA1 activity might restore genomic integrity and enhance neuronal resilience, while interventions targeting POLR3 C-mediated transcriptional dysregulation could mitigate protein aggregation and cellular stress in PD.

The investigation into the intersection of the immune system and PD is at the cutting edge of neuroscientific research, providing critical insights into the complex etiology of this neurodegenerative disease [45]. PD, classically characterized by the loss of dopaminergic neurons in the substantia nigra and the accumulation of α -synuclein, is increasingly being examined through the prism of immune dysregulation, a perspective informed by advances in neuroimmunology [18]. This rapidly evolving field has uncovered a complex interaction between innate and adaptive immune responses and PD pathogenesis [46]. A key finding is the role of

chronic neuroinflammation, characterized by microglial activation and peripheral immune cell infiltration, in the pathology of PD [47]. This inflammatory milieu is thought to exacerbate neuronal damage and hasten disease progression [48]. Additionally, α -synuclein aggregates, central to PD pathology, are believed to instigate immune responses, thereby amplifying neuroinflammation and contributing to neuronal degradation [49]. The role of the immune system in PD extends beyond neuroinflammation. Immunological factors, including cytokine profiles and the presence of autoantibodies, have been implicated in both the initiation and progression of PD [50]. In this research, CD4 memory-activated T cells, follicular helper T cells, gamma delta T cells, and resting dendritic cells were significantly enriched in the control group, suggesting a preserved immune surveillance and regulatory balance under physiological conditions. In contrast, the disease group exhibited heightened infiltration of regulatory T cells (Tregs), M0 macrophages, resting mast cells, activated mast cells, and neutrophils, reflecting a shift toward an immunosuppressive and proinflammatory microenvironment. The observed immune cell profile underscores the immunopathological shifts associated with disease progression. Elevated levels of Tregs and M0 macrophages suggest an immune evasion mechanism, where immunosuppressive signals may dampen anti-inflammatory responses, fostering a permissive environment for disease progression. Increased infiltration of mast cells and neutrophils highlights a heightened inflammatory state, which may exacerbate tissue damage and neuroinflammation. The depletion of CD4 memory-activated T cells, follicular helper T cells, and gamma delta T cells in the disease group points to impaired adaptive immune responses, potentially compromising immune surveillance and pathogen clearance. From a diagnostic perspective, these findings suggest that profiling immune cell populations could serve as a valuable biomarker strategy for early disease detection and monitoring. Therapeutically, modulating the immune landscape-such as enhancing T cell activation while suppressing the pro-inflammatory activity of mast cells and neutrophils-could represent a targeted approach to restoring immune balance and mitigating disease progression. These insights provide a foundation for developing immunomodulatory therapies aimed at recalibrating the immune microenvironment in disease contexts.

The recent surge of interest in the relationship between PD and metabolic processes represents a significant shift in modern medical research. With the emergence of advanced bioinformatics, there has been a groundbreaking expansion in our understanding of the molecular complexities of PD and its related pathologies [51–53].

This collective research effort is crucial in enhancing our understanding of the molecular mechanisms that underlie PD and its subsequent manifestations. Our study specifically addresses an important gap in the field, focusing on PMGs within the context of PD. Leveraging extensive datasets from the GEO (GSE6613 and GSE7621), we employed sophisticated analytical tools, including GO, KEGG, and GSEA. These methods have been pivotal in constructing a comprehensive predictive model that sheds light on the complex role of PMGs in PD pathogenesis. While our research establishes a fundamental theoretical framework, it also marks a stepping stone for future investigations into metabolic dysregulation in PD and the potential for therapeutic interventions targeting these disturbances. However, it is important to recognize that our study, despite its innovative approach, highlights the need for further empirical research to validate the primary mechanisms underlying PD. This crucial validation process should be pursued through extensive in vivo and in vitro studies, which are essential for deepening our understanding of PD and guiding the development of effective treatments.

Conclusions

In the complex pathobiology of PD, characterized by a multifaceted network of targets, pathways, signaling cascades, and regulatory mechanisms, PMGs emerge as pivotal players. Genes such as NME7, PKM, RRM2, POLR3 C, POLA1, PDE6 C, PDE9 A, PDE11 A, and AMPD1 are integral to the disease's molecular framework, orchestrating key metabolic and signaling processes. Among them, POLA1 and POLR3 C are particularly significant, exerting profound influence on metabolic regulation and cellular homeostasis in PD.

Abbreviations

GO	Gene Ontology
TCM	Traditional Chinese medicine
MF	Molecular functions
KEGG	Kyoto Encyclopedia of Genes and Genomes
GEO	Gene Expression Omnibus
PMGs	Purine Metabolism Genes
BP	Biological processes
CC	Cellular components
DEGs	Differentially Expressed Genes

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12883-025-04167-8.

Supplementary Material 1

Authors' contributions

The manuscript was written and corrected by Yao Wang and Dongchuan Wu. Data gathering was overseen by Yao Wang and Man Zheng. Man Zheng and Tiantian Yang conceived and designed this article, in charge of syntax

modification and revised of the manuscript. The final manuscript version has been read and approved by all of the writers.

Funding

This project was funded by the Natural Science Foundation of Shandong Province (ZR2022QH196).

Data availability

The datasets generated and/or analysed during the current study are available in the [GEO] repository. https://www.ncbi.nlm.nih.gov/geo/'

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Psychiatry, Shandong Daizhuang Hospital, Jining, China.
²Dongying City Traditional Chinese Medicine Hospital, Dongying, People's Republic of China.
³Dongying People'S Hospital (Dongying Hospital of Shandong Provincial Hospital Group), Dongying, People's Republic of China.
⁴Department of Traditional Chinese Medicine, Shandong Provincial Hospital-affiliated to, Shandong First Medical Universityaq, Jinan, China.

Received: 13 February 2025 Accepted: 1 April 2025 Published online: 16 April 2025

References

- 1. Dickson DW. Neuropathology of Parkinson disease. Parkinsonism Relat Disord. 2018;46 Suppl 1(Suppl 1):S30–3.
- Grimes D, Fitzpatrick M, Gordon J, Miyasaki J, Fon EA, Schlossmacher M, Suchowersky O, Rajput A, Lafontaine AL, Mestre T, et al. Canadian guideline for Parkinson disease. CMAJ. 2019;191(36):E989–1004.
- 3. Singer C, Reich SG. Parkinson Disease. Clin Geriatr Med. 2020;36(1):xiii–xiv.
- Zesiewicz TA. Parkinson Disease. Continuum (Minneap Minn). 2019;25(4):896–918.
- Marshall K, Hale D. Parkinson Disease. Home Healthc Now. 2020;38(1):48–9.
- Halli-Tierney AD, Luker J, Carroll DG. Parkinson Disease. AM FAM PHYSI-CIAN. 2020;102(11):679–91.
- Xiao B, Tan EK. Sniffing for Parkinson Disease. AM J MED. 2023;136(5):411–2.
- Bloem BR, Okun MS, Klein C. Parkinson's disease. Lancet. 2021;397(10291):2284–303.
- Pinnell JR, Cui M, Tieu K. Exosomes in Parkinson disease. J NEUROCHEM. 2021;157(3):413–28.
- 10. Semenova EA, Hall E, Ahmetov II. Genes and Athletic Performance: The 2023 Update. Genes (Basel) 2023;14(6):12–35.
- Jo DH, Kim JH, Kim JH. Tumor Environment of Retinoblastoma. Intraocular Cancer ADV EXP MED BIOL. 2020;1296:349–58.
- Zhang Y, Zhang Z. The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. CELL MOL IMMUNOL. 2020;17(8):807–21.
- Yang S, Zhang B, Tan W, Qi L, Ma X, Wang X. A Novel Purine and Uric Metabolism Signature Predicting the Prognosis of Hepatocellular Carcinoma. FRONT GENET. 2022;13: 942267.
- 14. Liu J, Hong S, Yang J, Zhang X, Wang Y, Wang H, Peng J, Hong L. Targeting purine metabolism in ovarian cancer. J OVARIAN RES. 2022;15(1):93.
- Yin J, Ren W, Huang X, Deng J, Li T, Yin Y. Potential Mechanisms Connecting Purine Metabolism and Cancer Therapy. FRONT IMMUNOL. 2018;9:1697.

- Shatova OP, Butenko EV, Khomutov EV, Kaplun DS, Sedakov IE, Zinkovych II. Metformin impact on purine metabolism in breast cancer. Biomed Khim. 2016;62(3):302–5.
- Chen X, Chen J. miR-10b-5p-mediated upregulation of PIEZO1 predicts poor prognosis and links to purine metabolism in breast cancer. Genomics. 2022;114(3): 110351.
- Zhao S, Zhang L, Ji W, Shi Y, Lai G, Chi H, Huang W, Cheng C. Machine learning-based characterization of cuprotosis-related biomarkers and immune infiltration in Parkinson's disease. FRONT GENET. 2022;13:1010361.
- Xu J, Li J, Sun YJ, Quan W, Liu L, Zhang QH, Qin YD, Pei XC, Su H, Chen JJ. Identification of key genes and signaling pathways associated with dementia with Lewy bodies and Parkinson's disease dementia using bioinformatics. FRONT NEUROL. 2023;14:1029370.
- Liu L, Cui Y, Chang YZ, Yu P. Ferroptosis-related factors in the substantia nigra are associated with Parkinson's disease. Sci Rep. 2023;13(1):15365.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. NUCLEIC ACIDS RES. 2015;43(7): e47.
- 22. Hackenberger BK. R software: unfriendly but probably the best. Croat Med J. 2020;61(1):66-68.1.
- Gustavsson EK, Zhang D, Reynolds RH, Garcia-Ruiz S, Ryten M. ggtranscript: an R package for the visualization and interpretation of transcript isoforms using ggplot2. Bioinformatics. 2022;38(15):3844–6.
- Xu D, Chu M, Chen Y, Fang Y, Wang J, Zhang X, Xu F. Identification and verification of ferroptosis-related genes in the pathology of epilepsy: insights from CIBERSORT algorithm analysis. FRONT NEUROL. 2023;14:1275606.
- Darweesh S, Raphael KG, Brundin P, Matthews H, Wyse RK, Chen H, Bloem BR. Parkinson Matters. J Parkinsons Dis. 2018;8(4):495–8.
- 26. Reich SG, Savitt JM. Parkinson's Disease. Med Clin North Am. 2019;103(2):337–50.
- 27. Balestrino R, Schapira A. Parkinson disease. EUR J NEUROL. 2020;27(1):27–42.
- 28. Homayoun H. Parkinson Disease. ANN INTERN MED. 2018;169(5):C33-48.
- 29. Kim SD, Allen NE, Canning CG, Fung V. Parkinson disease. Handb Clin Neurol. 2018;159:173–93.
- 30. Qi G, Mi Y, Shi X, Gu H, Brinton RD, Yin F. ApoE4 Impairs Neuron-Astrocyte Coupling of Fatty Acid Metabolism. CELL REP. 2021;34(1): 108572.
- Yang K, Li J, Tao L. Purine metabolism in the development of osteoporosis. BIOMED PHARMACOTHER. 2022;155: 113784.
- Furuhashi M. New insights into purine metabolism in metabolic diseases: role of xanthine oxidoreductase activity. Am J Physiol Endocrinol Metab. 2020;319(5):E827–34.
- Yu F, Quan F, Xu J, Zhang Y, Xie Y, Zhang J, Lan Y, Yuan H, Zhang H, Cheng S, et al. Breast cancer prognosis signature: linking risk stratification to disease subtypes. BRIEF BIOINFORM. 2019;20(6):2130–40.
- Jiang Z, Shen H, Tang B, Yu Q, Ji X, Wang L. Quantitative proteomic analysis reveals that proteins required for fatty acid metabolism may serve as diagnostic markers for gastric cancer. CLIN CHIM ACTA. 2017;464:148–54.
- Lei C, Zhongyan Z, Wenting S, Jing Z, Liyun Q, Hongyi H, Juntao Y, Qing Y. Identification of necroptosis-related genes in Parkinson's disease by integrated bioinformatics analysis and experimental validation. Front Neurosci. 2023;17:1097293.
- Georgiou M, Robson AG, Singh N, Pontikos N, Kane T, Hirji N, Ripamonti C, Rotsos T, Dubra A, Kalitzeos A, et al. Deep Phenotyping of PDE6C-Associated Achromatopsia. Invest Ophthalmol Vis Sci. 2019;60(15):5112–23.
- Starokadomskyy P, Escala PA, Burstein E. Immune Dysfunction in Mendelian Disorders of POLA1 Deficiency. J CLIN IMMUNOL. 2021;41(2):285–93.
- Dallavalle S, Musso L, Cincinelli R, Darwiche N, Gervasoni S, Vistoli G, Guglielmi MB, La Porta I, Pizzulo M, Modica E, et al. Antitumor activity of novel POLA1-HDAC11 dual inhibitors. EUR J MED CHEM. 2022;228: 113971.
- Bellelli R, Belan O, Pye VE, Clement C, Maslen SL, Skehel JM, Cherepanov P, Almouzni G, Boulton SJ. POLE3-POLE4 Is a Histone H3–H4 Chaperone that Maintains Chromatin Integrity during DNA Replication. MOL CELL. 2018;72(1):112–26.
- Cincinelli R, Musso L, Guglielmi MB, La Porta I, Fucci A, Luca DE, Cardile F, Colelli F, Signorino G, Darwiche N, et al. Novel adamantyl retinoid-related molecules with POLA1 inhibitory activity. BIOORG CHEM. 2020;104: 104253.
- 41. Li YF, Hou QQ, Zhao S, Chen X, Tang M, Li L. Identification of tumorspecific neoantigens and immune clusters of hepatocellular

carcinoma for mRNA vaccine development. J Cancer Res Clin Oncol. 2023;149(2):623–37.

- 42. Lata E, Choquet K, Sagliocco F, Brais B, Bernard G, Teichmann M. RNA Polymerase III Subunit Mutations in Genetic Diseases. Front Mol Biosci. 2021;8: 696438.
- Harich B, van der Voet M, Klein M, Cizek P, Fenckova M, Schenck A, Franke B. From Rare Copy Number Variants to Biological Processes in ADHD. Am J Psychiatry. 2020;177(9):855–66.
- 44. Saadat A, Gouttenoire J, Ripellino P, Semela D, Amar S, Frey BM, Fontana S, Mdawar-Bailly E, Moradpour D, Fellay J, et al. HEV Human Genetics Collaborators. Inborn errors of type I interferon immunity in patients with symptomatic acute hepatitis E. Hepatology. 2024;79(6):1421-31. https://doi.org/10.1097/HEP.00000000000701.
- Lauritsen J, Romero-Ramos M. The systemic immune response in Parkinson's disease: focus on the peripheral immune component. TRENDS NEUROSCI. 2023;46(10):863–78.
- Tan EK, Chao YX, West A, Chan LL, Poewe W, Jankovic J. Parkinson disease and the immune system - associations, mechanisms and therapeutics. NAT REV NEUROL. 2020;16(6):303–18.
- Tansey MG, Wallings RL, Houser MC, Herrick MK, Keating CE, Joers V. Inflammation and immune dysfunction in Parkinson disease. NAT REV IMMUNOL. 2022;22(11):657–73.
- Munoz-Delgado L, Macias-Garcia D, Perinan MT, Jesus S, Adarmes-Gomez AD, Bonilla TM, Buiza RD, Jimenez-Jaraba M, Benitez ZB, Diaz BR, et al. Peripheral inflammatory immune response differs among sporadic and familial Parkinson's disease. NPJ Parkinsons Dis. 2023;9(1):12.
- 49. Abdi IY, Ghanem SS, El-Agnaf OM. Immune-related biomarkers for Parkinson's disease. NEUROBIOL DIS. 2022;170: 105771.
- Schonhoff AM, Williams GP, Wallen ZD, Standaert DG, Harms AS. Innate and adaptive immune responses in Parkinson's disease. PROG BRAIN RES. 2020;252:169–216.
- Wang M, Li T, Gao R, Zhang Y, Han Y. Identifying the potential genes in alpha synuclein driving ferroptosis of Parkinson's disease. Sci Rep. 2023;13(1):16893.
- Wang H, Dou S, Wang C, Gao W, Cheng B, Yan F. Identification and Experimental Validation of Parkinson's Disease with Major Depressive Disorder Common Genes. MOL NEUROBIOL. 2023;60(10):6092–108.
- Wu Z, Cai Z, Shi H, Huang X, Cai M, Yuan K, Huang P, Shi G, Yan T, Li Z. Effective biomarkers and therapeutic targets of nerve-immunity interaction in the treatment of depression: an integrated investigation of the miRNA-mRNA regulatory networks. Aging (Albany NY). 2022;14(8):3569–96.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.