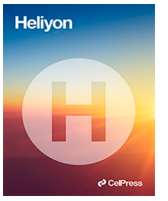
Heliyon 10 (2024) e38403

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Utilizing an integrated bioinformatics and machine learning approach to uncover biomarkers linking ulcerative colitis to purine metabolism-related genes

Tian Chen [1], Yiqiu Tao [1], Qingyuan Wang [1], Yanni Pei, Zhenhua Zhao, Wei Yang [**], Yafeng Lu [*]

Department of Anorectology, Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, China

ARTICLE INFO

Keywords:
Ulcerative colitis
Purine metabolism genes
Lasso regression
SVM-RFE
Bioinformatics

ABSTRACT

Background: Ulcerative colitis (UC) is an increasing incidence of inflammatory disorder in the colon mucosa. One of the current research focuses is the alteration of metabolic networks in UC. One of the important aspects of this metabolic shift is the expression of purine metabolism genes (PMGs) vital for nucleic acid synthesis. Nevertheless, the precise function of PMGs in the pathophysiology of UC is not yet fully known.

Methods: To this end, this study used state-of-the-art bioinformatics tools and approaches to discover and confirm the PMGs involved in UC. All the 114 candidate PMGs were compared for their expression levels. GSEA and GSVA were applied to define the functional and pathway implications of these PMGs. Lasso regression and SVM-RFE approaches were used for the identification of hub genes and to assess the diagnostic potential of eight PMGs in UC classification. The relationship between these critical PMGs and clinical features was also systematically evaluated as well. The expression levels of these eight PMGs were validated using datasets GSE206285 and GSE179285.

Results: Using bioinformatics and machine learning, this work seeks to establish the involvement of PMGs in UC. From the LASSO and SVM models, 114 DE PMGs were selected and investigated to build a stable predictive model. Based on these studies, the following genes: IMPDH1, GUK1, POLE3, ADCY3, ADCY4, PDE6B, PNPT1 and PDE4D were suggested as potential biomarkers of UC. Gene ontology enrichment analysis revealed that these genes are implicated in the biological processes of particular relevance to immune and inflammatory responses. The study also provided a lot of information on the interaction between immune cells and PMGs indicating that these genes may control some immune-related pathways in UC. Moreover, drug-gene interaction analysis presents potential therapeutic opportunities for potential drug targets which were further confirmed through molecular docking. Mendelian randomization analysis revealed that ADCY4 and PDAZN are involved in PMG-related processes, thus opening new possibilities for treatment.

* Corresponding author.
** Corresponding author.
E-mail addresses: yangweiyishi@163.com (W. Yang), luyafeng2023@126.com (Y. Lu).
[1] Co-first author.

https://doi.org/10.1016/j.heliyon.2024.e38403
Received 14 February 2024; Received in revised form 19 September 2024; Accepted 24 September 2024
Available online 24 September 2024

Conclusions: This work reveals eight PMGs closely related to UC and provides new perspectives on possible markers of this inflammatory disease. These findings not only increase the understanding of the pathogenesis of UC but also offer potential for improving the surveillance of disease and its progression.

1. Introduction

Actually, ulcerative colitis (UC), an idiopathic inflammatory bowel disease (IBD) with the cyclic remitting-relapsing mode of the disease, is recognized as a large-scale public health problem on the international level [1]. Associated with a number of manifestations including abdominal pain, pus and mucus discharge, and bloody diarrhea, UC is predominantly seen in developed countries, though its incidence is increasing in developing countries [2]. This is because the ubiquity of this disease is due to the fact that it does not discriminate against the persons it affects; people of all ages can be affected. The cause of UC is still not fully understood and there are numerous factors that have been implicated; Clinically, UC is delineated into two distinct phases: It is divided into the active phase in which there are symptoms, and the remission phase in which there are no symptoms [3]. The authors have also noted that UC is characterized as a progressive disease that leads to severe intestinal complications including but not limited to colorectal cancer which has a direct effect on the life of the patients [4]. The diagnosis of UC, a challenging task demands total clinical evaluation, laboratory investigation, imaging, endoscopy, and histopathological examination Sometimes even the most experienced clinicians may find it challenging to come up with a diagnosis of UC. The treatment of UC has been characterized by the employment of 5-aminosalicylate, corticosteroids, and immunomodulators as the main therapeutic options [5]. Biological therapies and small molecular compounds have come into practice starting in the twenty-first century and are particularly useful in controlling moderate to severe UC [6]. However, the overall and therapeutic outcomes of the present regimens are still unsatisfactory; only about 40 % of patients experience clinical remission each year; and, a significant number (20–30 %) of patients may necessitate surgical intervention [7]. Against this background, the importance of current literature in the discovery of the elements of immunity-inflammation in UC's pathogenesis cannot be overemphasized [8]. Such efforts are crucial when trying to establish accurate diagnostic biomarkers and/or design targeted eventual therapies. In this connection, the identification of molecular markers of UC at an early stage is considered the basis for intervention measures [9]. This approach of pre-emptive intervention, that is, before the onset of clinical symptoms is expected to transform the management and outcome of UC. Therefore, this strategy is expected to shift the treatment paradigm in UC towards early diagnosis and treatment with the aim of improving the disease prognosis.

In the context of cancer, metabolic reprogramming can be considered one of the hallmarks of cancer as it is fundamental to supporting cancer cell growth and survival. This reprogramming results in a specific metabolic profile that is manifest in tumor cells and consequently alters the TME. The TME which is characterized by various cell types and conditions such as low oxygen and nutrient

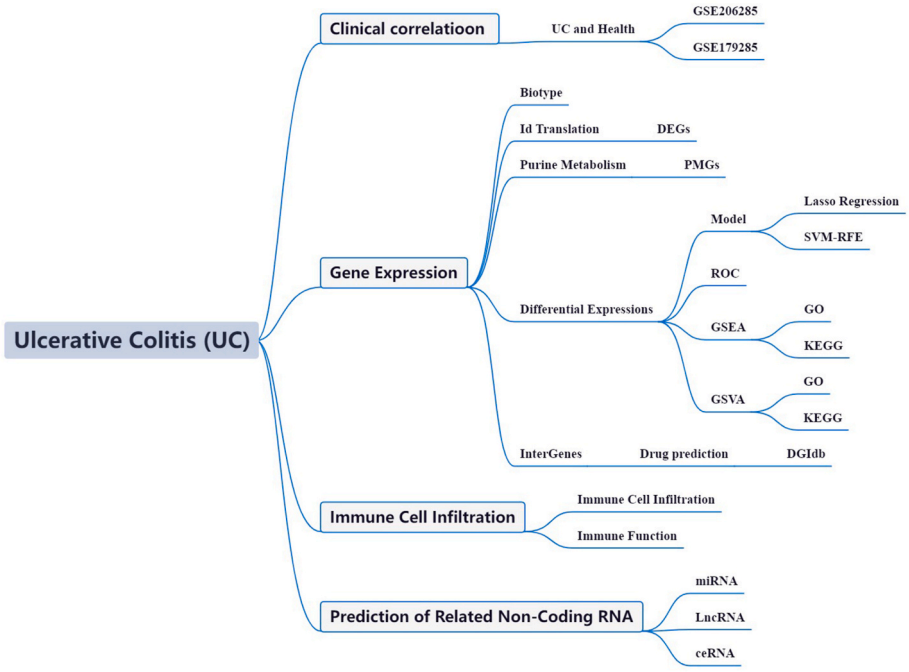


Fig. 1. Framework.

supply is a key factor in the advancement of cancer [10]. This has shifted focus towards non-tumoral immune infiltration within the TME as the most recent developments have shown. Recent findings indicate that immune response is accompanied by profound changes in tissue metabolism where nutrient availability is low, oxygen consumption is increased, and the production of reactive species is enhanced [11]. This metabolic rewiring has profound consequences for immune cell function and development; these provide several avenues for harnessing metabolism as a means of improving immune-based therapies [12]. It is required since, they are the building blocks of DNA and RNA molecules and are comprised of other biomolecules such as ATP and NADH besides being involved in certain functions or processes in the cell like energy production, signal transmission, synthesis of fatty acids [13]. Their roles also involve regulating immune responses as well as host-pathogen interactions [14]. Mammalian cells are not able to synthesize purines de novo to the same extent as pyrimidines and their requirements are mostly provided through salvage pathways. However, for cells that divide at a very fast rate for example cancer cells, the demand for purines is augmented by increasing de novo synthesis [15]. Purine antimetabolites were among the first anticancer agents and are still used in the treatment of different leukemias and some non-neoplastic diseases operating through inhibition of DNA synthesis and cell proliferation [16]. The identification of ribosomes, structures that are intimately associated with the cell cycle has revealed a new target for therapy within the metabolism of purines [17]. This work is likely to benefit from the integration of purine metabolic strategies with immunotherapy especially in the treatment of UC. Nevertheless, the place of purine metabolism in immunogenicity and the efficiency of immunotherapy remains a field that is relatively not well explored. However, there is a gap in the identification and analysis of the existing PMGs with immunotherapy for UC in the current study, this systematic review aims to try to fill this gap. Thus, it tries to change the current perception and management of UC and, thus, open new horizons in the treatment of the disease.

The UC Initiative is a novel approach that integrates high-throughput RNA-seq data and detailed clinical phenotypes to map a new avenue into the transcriptional and molecular complexity of UC [18–20]. The application of bioinformatics to these large and comprehensive datasets has provided valuable information about the UC pathogenesis. However, there is a major gap that has not been fully exploited in using bioinformatics to uncover the relationship of PMGs to UC. In doing so, this study endeavors to fill the research gap, using UC-associated GEO data sets to elucidate the role and consequences of PMGs in the development of UC, as illustrated in Fig. 1. This research study aims to create a better framework for the study of molecular mechanisms of UC so as to allow for new and effective therapeutic protocols to be developed. This work not only expands the current understanding of UC by focusing on PMGs but also suggests that it is possible to develop new targeted therapies for this multisystem disease.

In an endeavor to systematically define UC and gain preliminary insights, we embarked on an exploratory study utilizing patient-derived gene expression profiles from GEO databases. Therefore, the GSE206285 set was set as the main cohort, while the GSE179285 set was set as a validation cohort. In the next step, applying a strict matching method of PMGs, we conducted differential expression studies and constructed prognostic risk models. The current study therefore offered a set of PMGs, which yielded the potential for use of the markers in UC patients and endorsed the feasibility of the exploration of the proposed methodological framework. Therefore to gain insight into the functions of the identified genes, we carried out various analyses that include GO, KEGG, and GSEA. These comprehensive analyses were further enriched with the interaction with several databases that indicated a more detailed picture of the implicated PMGs in the cell context and molecular interactions and regulation. We also applied the methods to depict the immunological features and alterations by analyzing the immune cell infiltration and the functional and transcriptional alterations. This therefore not only adds to the understanding of what the implications of PMGs in UC entail but also lays a valuable groundwork for future-orientated treatment paradigms in the management of this still baffling neurodegenerative disorder.

2. Material and methods

The methods described by Zi-Xuan Wu et al. in 2023 were used in this work [21]. We used the methods of Zixuan Wu et al. 2023 [21] 2024 [22].

2.1. Raw data

The gene expression omnibus datasets GSE206285 and GSE179285 were used in this study. We used two platforms for this study that were GPL13158 and GPL6480. GSE206285 was used as the training set while GSE179285 was used as the test set. We also obtained PMGs from the Molecular Signatures Database (MSigDB) (Table S1).

2.2. Delineation of differentially expressed genes (DEGs)

Thus, to achieve the required mRNA profiles with high specificity, we improved the strategy based on Perl scripts for cross-analyzing the transcriptional data within GSE58331. Following normalization procedures, we applied stringent criteria for identifying differential expressions among PMGs: FDR is < 0.05 and $|\log2FC| \geq 1$. This procedure was quite stringent which aided in the identification of various PMGs that were differentially regulated for further characterization. To study the reciprocal relations of these genes Pearson's correlation coefficient was applied and for comprehensive correlation analysis corrplot package in R was used. This was important in ordering the given genes of interest with respect to their statistical association with the identified modules.

To build a prognostic model for PMGs we used the glmnet and survival packages. The predictor signature for PMG was established and then validated through the using LASSO penalized Cox regression and Uni-variate Cox regression analysis. The risk score for each UC patient was determined based on the formula: Finally, the gene expression model is given as (Co-efficient of DEGs1 $\times$ expression of DEGs1) + (Co-efficient of DEGs2 $\times$ expression of DEGs2) + … + (Co-efficient of DEGsn $\times$ expression of DEGsn). This risk score was

then used to stratify patients into two subgroups: Low risk was categorized as having less than the requisite number of risk factors considered to be the median and high risk was categorized as having equal to or more than the requisite number of risk factors considered to be the median.

In lasso regression, the low and high risk of patients was deemed and so, the relative plots were created. Subsequently, we carried out the confidence interval test and risk ratio in addition to the heatmap package for creating the forest diagram. The curves of survival were then plotted in order to study the probability of the high-risk and low-risk categories of the patients.

2.3. Functional enrichment analysis: GO and KEGG pathways

To unravel biological meanings and function of DEGs in the pathway, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis. We also investigated how these PMGs with dysregulated expression implicated the BP, MF, and CC using R This study aimed at establishing and characterize the major biological processes and molecular mechanisms involving these genes to gain more understanding of the role of the genes in disease states and on possible therapeutic interventions. Aside from just making the classification of the DEGs, we wanted to unveil the likely correlation between purine metabolism and other processes and diseases.

2.4. The process of model construction and immune cell infiltration

In order to get the closest to the ideal model we employed the Lasso regression analysis with the help of the glmnet tool and cross-validation. This approach was quite effective in minimizing overfitting than the previous approaches thus enhancing the model's capability in handling biological data complexities. We finally used SVM-RFE with a powerful e1071 package and created a machine-learning model very cautiously. The other critical process that was adopted in the development of the model was cross-validation which helped in assessing the error rate of the model as well as the general effectiveness of the entire process. In the present analysis, to predict the enrolment rate, the algorithm i. e. , the random forest which is developed through the ensemble learning method was employed. This was done by making several decision trees and then using their results combined in a way that would minimize over fitting the data and maximize the performance of the model. Hence, the random feature selection and bootstrap sampling methods that are used in this method improved the diversities of the decision trees and thus improved the performance of the model. In this case, the focus was made on the differentially expressed genes which can be drawn from the randomForest package using the usual steps in R, and the ggplot2 package was used in the same manner as before. For the grading of these feature genes in the last step, we have considered Lasso regression, Random Forest, and SVM models that gave us the overall picture of their involvement in the disease process. In addition, based on CIBERSORT, we not only obtain the proportion of the immune cells but also know more about the immune context associated with the disease. Apart from enhancing the reliability of the disease categorization, this broad and strong approach also offered unique perspectives on the molecular etiology of the disease.

2.5. Gene set enrichment and variation analyses

Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA) were used to identify functional changes and pathway differences across a number of samples. These analyses allowed for the grouping of gene sets and pathways and the use of scores and graphs to effectively determine the most biologically active processes and pathways in the various risk stratifications. Here R was used to analyze the impact of the DE PMGs on BP, MF, and CC as well as on the pathways which gave a better insight into the roles of the DE PMGs in the disease process.

2.6. Drug-gene interaction insights

Due to modern bioinformatics at present, biological models were created, and biomarkers were used in the diagnosis of diseases. On the other hand, the transition from the biomarker discovery phase to the use of biomarkers in the clinic is still problematic. Therefore, the use of those markers for the prediction of therapeutic interventions in UC is likely to have a great effect on the future place of UC in the system of its treatment in the future. Appreciating that such biomarkers form the basis for numerous treatments, drug prediction represents a critical focus. Therefore, using the obtained hub genes, we applied, for the first time to our knowledge, the DGIdb approach to predict possible drug interactions with them, providing a basis for selective therapy.

2.7. The relation between common miRNAs and LncRNAs

The regulation of the genetic processes, however, occurs by way of the non-coding RNA transcripts that include the microRNAs (miRNAs), and long non-coding RNAs (lncRNAs). Non-coding RNAs can also affect gene expression by degrading the mRNA and their translation, however, miRNAs are involved in both degradation of the mRNA and suppression of their translation whereas lncRNAs are more than 200 nucleotides in length and are involved in several other cellular processes such as chromosome remodeling, transcription, and interference. Such studies indicate that various lncRNAs and miRNAs have a rather complex interaction and that there is an intricate interplay between the systems containing these agents and other elements. This interaction has raised awareness of competitive endogenous RNAs (ceRNAs); the possibility of a lncRNA regulating the expression of a gene through possessing a miRNA has been observed. Gathering these data, the current work focused on possible commonalities of miRNAs and lncRNAs regulatory

mechanisms and developmental pathways in UC with the purpose of identifying novel diagnostic and therapeutic targets for this complex disease.

2.8. Construction of an mRNA-miRNA-LncRNA network

To establish the wiring diagram of mRNA, miRNA, and lncRNA cross-talk in UC, we downloaded the target genes from miRTarBase and PrognoScan databases because both of these databases have confirmed miRNA-lncRNA-target gene relationships. The UC-associated genes were then extracted according to the targeted genes of common mRNA-miRNA-lncRNA and then displayed as the regulated network based on Cytoscape to elucidate the overall molecular relationship of the disease.

2.9. Mendelian randomization analysis

For our main analysis, we used the TwoSampleMR package in R to perform an association analysis between our exposure and outcome variables in our GWAS summary data. Here, ADCY4 and PNPT1-related expressions were used as exposure, and UC as the outcome of interest. The analysis entailed: 1. Instrumental Variables (IVs) Configuration: For the identification of the strongly associated exposures, the P-value threshold of $<5 \times 10^{-8}$ was used for ADCY4 and PNPT1-related expressions. We restricted the analysis to the European population with imputation to r2 > 0.30, minor allele frequency >0.01, and within 10,000 kb to avoid LD and reduce pleiotropic effects. 3. Statistical Strength Configuration: To ensure that the validity of instrumental variables the F-statistic was calculated ($F = \beta^2/SE^2$) where those with $F < 10$ were deemed not strong enough to overcome confounding influence.

From the given GWAS data, we have been able to point out the SNPs that are associated with the IVs. Using the 'harmonise_data' function in the TwoSampleMR software, we then harmonized the allelic direction of exposure and outcome and removed SNPs that were not compatible with this. The IVW approach to causal estimation was employed to investigate causal relationships and to further the understanding of the genetic basis of diseases with the variance of IVs being used as weights.

2.10. Active components-targets docking

The interactions of ADCY4 and PNPT1 were validated for the principal components and the targets of prediction. The protein configurations of the core targets were obtained from the Uniprot (https:The UniProt database was searched with the minimum resolution (Resolution) and source (Method) of protein X-ray with the condition that the protein structures of these configurations were retrieved from the RCSB PDB database (https://www. rcsb. org/). 2D structures of 6 active components of core targets were obtained from PubChen database (https:The 2D structures of these compounds were retrieved from PubChem database (https://pubchem. ncbi. nlm. nih. gov/) and these 2D structures were energy minimized using chem3d software. The binding affinities and

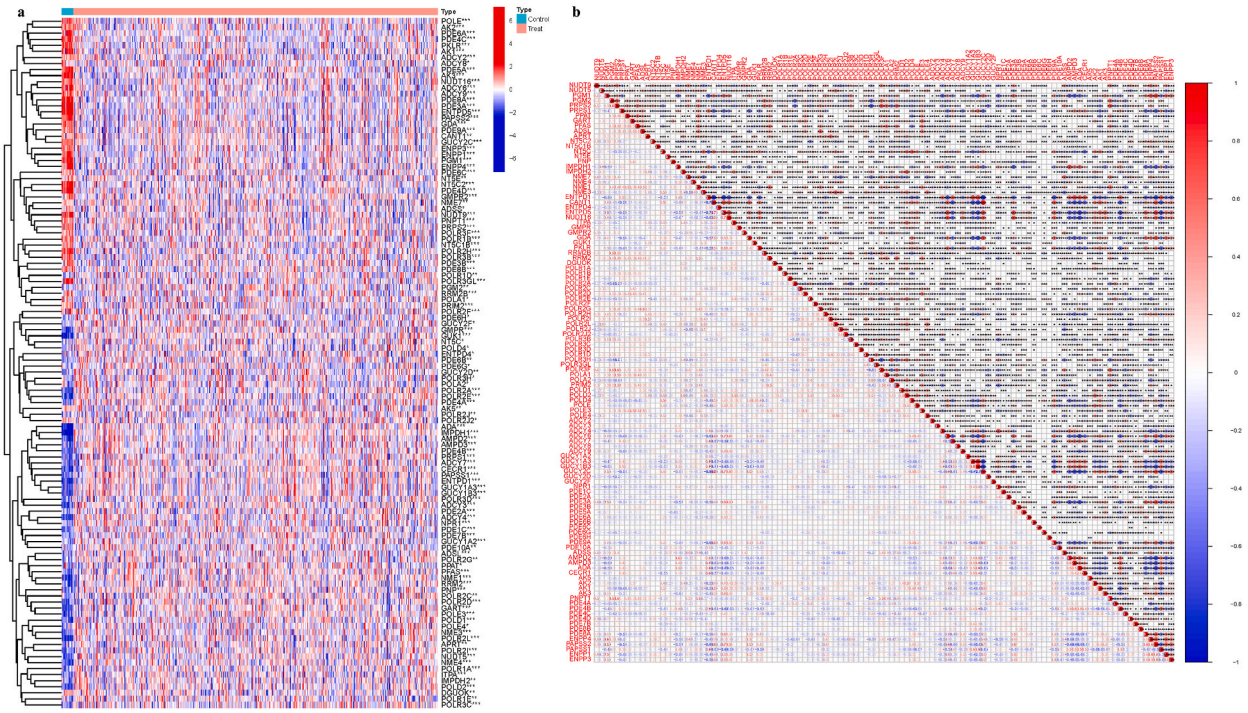


Fig. 2. Principal component analysis. (a) Analysis of difference. (b) Analysis of correlation.

the binding activities of active components and targets were performed by using SYBYL2. No software, and the active components of binding TotalScore greater than 3 were selected for sub-docking.

The crystal was then docked into the Pymol 2.4 Software (http://www. pymol. org/2/) and the structure was later export to AutoDockTools 1.5.6 for building the docking grid box for every target. It was done by Vina 1.1.2 software and the docking was successfully done and the molecules with the lowest binding energy in the docking conformation were chosen to compare the binding effect with the original ligands and the intermolecular interactions including hydrophobic, cation-π, anion-π, π-π stacking, hydrogen bonding and so on. Finally, the Pymol2. Molecular docking was done using the 4 software to analyze the output.

3. Results

3.1. Data about DEGs and principal component analysis

From the 114 assessed PMGs, we can see that some of them had more significant or less expression compared to the other PMGs. Furthermore, the gene clustering profile revealed that there are different clusters in the treatment and the control groups. That of the treatment groups are PDE7B, GUCY1A2, PDE10A, ADSL, POLR2G, PPAT, PFAS, NME1, RRM2, PNP, POLR2C, and POLR2D as obtained in Fig. 2a while the control Group observed included PDE8A, PDE3A, ENTPD5, PAPSS2 These PMGs were further analyzed for correlation and the correlation matrix is as shown in Fig. 2b (Table S2).

3.2. Enrichment analysis of PMGs

Finally, the GO enrichment analysis (Fig. 3a) indicated that there were 440 potential core target genes that participated in GO BP, MF, and CC. The MF category primarily involved guanyl nucleotide binding (GO: Among the biological process, the protein is involved

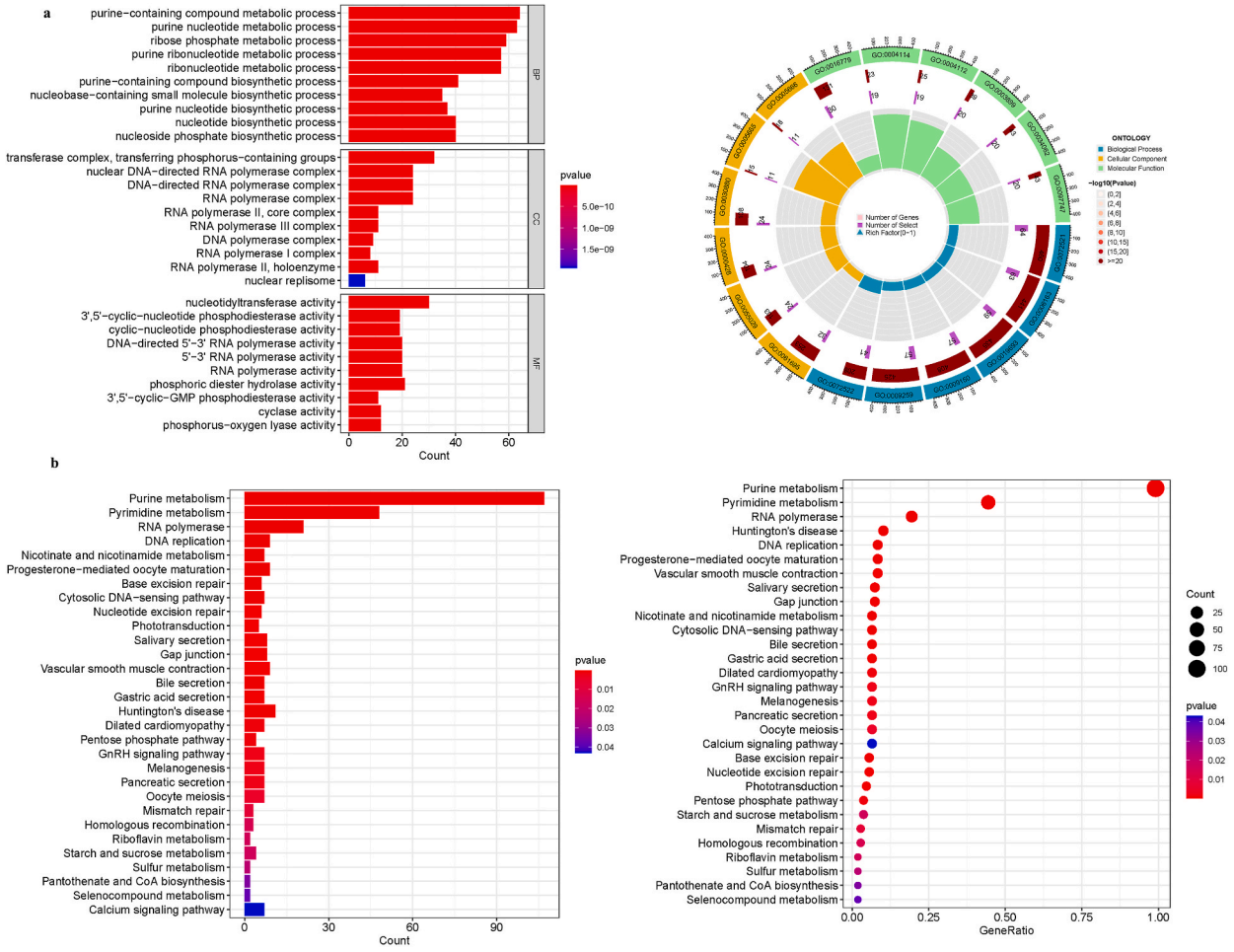


Fig. 3. For PMGs, GO and KEGG analyses were performed. (a): The GO circle illustrates the scatter map of the selected gene's logFC. (b): The KEGG barplot and bubble illustrates the scatter map of the logFC of the indicated gene.

in dGTP binding (GO:0019001), guanyl ribonucleotide binding (GO:0032561) and nucleoside binding (GO:0001882). The CC category was mainly associated with cell projection membrane (GO: Some of these are; Golgi apparatus (GO:0006514), vesicle lumen (GO:0031983), and cytoplasmic vesicle lumen (GO:0060205). The BP category included neutrophil-mediated immunity (GO: Neutrophils were enriched for granulocyte activation (GO:0002446), neutrophil activation involved in immune response (GO:0002283), and neutrophil degranulation (GO:0043312). From the upregulated differentially expressed genes, the KEGG enrichment analysis (Fig. 3b) recommended that they converged with Huntington's disease (KEGG hsa05016), the Calcium signaling pathway (KEGG hsa04020), and Purine metabolism (KEGG hsa00230) (Tables S3a–b).

3.3. Model construction

In the current study, the highest number of genes was chosen using LASSO and Cox regression analysis and we ensured that we chose the best threshold as shown in Fig. 4a and b. For the purpose of confirming the accuracy of our model and its ability to generalize, we developed an SVM-RFE machine-learning model. This model provided a high degree of accuracy with a score of 0.996 and the error rate was set to around 0. Fig. 4c and d: 00351 as seen in Fig. 4c and d. Moreover, the integration of the eight PMGs identified in this study by both LASSO and SVM-RFE analysis was shown to be quite comparable with a high degree of concordance thus supporting the prognosis model as depicted in Fig. 4e. Upon evaluating the model in relation to the 8 hub genes, we observed notably high accuracy rates for each gene: IMPDH1 (AUC = 0.999), GUK1 (AUC = 0.967), POLE3 (AUC = 0.887), ADCY3 (AUC = 0.929), ADCY4 (AUC = 0.960), PDE6B (AUC = 0.725), PNPT1 (AUC = 0.813) and PDE4D (AUC = 0.944) (Fig. 4f). Interestingly, the AUC can be one. With a similar trend in dataset GSE206285, the AUC of 1.000 (95 % CI 1.000–1.000) was also obtained, which supports the high reliability of the proposed prediction model (Fig. 4g) (Table 1 and S4). For the performance assessment of our work particularly the AUC, a close scrutiny was made on Fig. 4 to determine its AUC and found to be 0.996 which seemed quite good to us since it depicts how well we have done on the given predictions. To solve such problems, related to the lower AUC value of some gen, it is necessary to take into account genetic factors that affect the outcome. However, it is worth underlining that the AUC values for the above genes are, in general, close to the significance level of 0.7. This combined result significantly increases the credibility, reliability, and accuracy of our proposed predictive model reducing any limitations that may exist in its clinical and scientific use.

3.4. Gene set enrichment analysis

In this study, we employed the AUC and Rank to indicate the order of each gene in the test group as well as the validation results of the test group. Based on these results, we supposed that ADCY4 and PNPT1 could be the closest genes with which the disease is linked. Therefore having reviewed the literature and analyzed the sensitivity of the model hub gene; it was deduced that; ADCY4 and PNPT1 could be dent as the most important genes in UC. For GO analysis, it was found that ADCY4 had an association with BP ameboidal type cell migration, BP b cell activation, and BP cell chemotaxis. On the other hand, PNPT1 was enriched in the biological process of BP cellular response to biotic stimulus, BP cellular response to molecule of bacterial origin, and BP positive regulation of inflammatory response (Fig. 5a). The target genes of ADCY4 were mainly involved in ECM receptor interaction, complement and coagulation

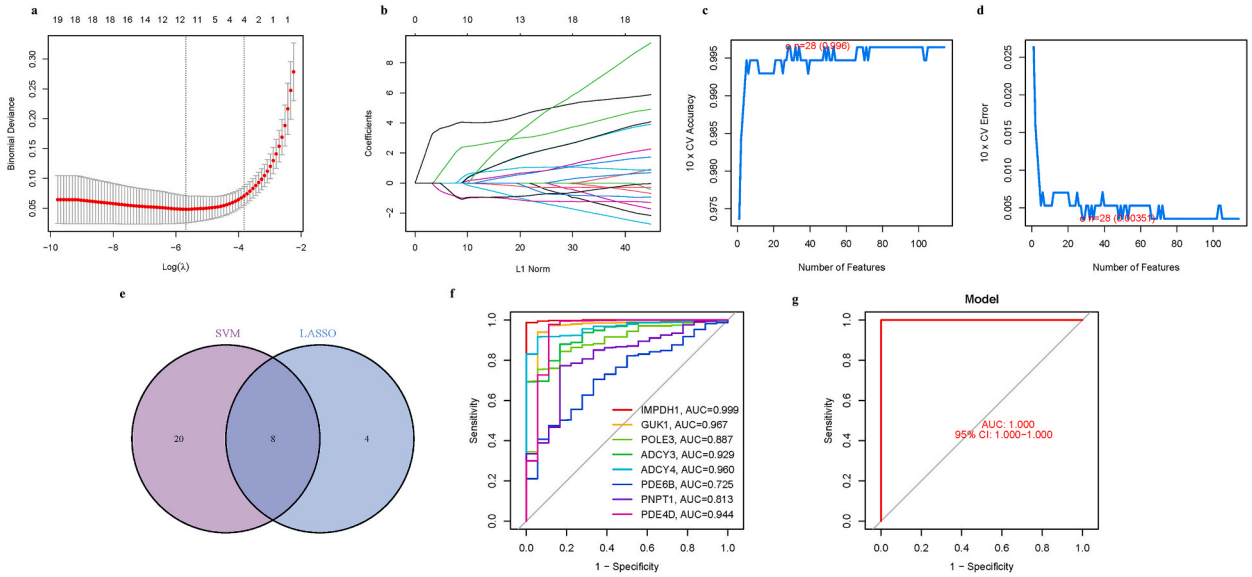


Fig. 4. The evolution of the signature of PMGs. (a): UC- LASSO regression analysis of the 8 UC related genes. (b): The use of cross-validation was made to determine optimal parameters for the LASSO regression model. (c–d): This is where we are going to look at how accurate this model is and the possible errors it may entail. (e): Venn. (f): AUC of the eight hub genes: (g): Train group's AUC of the model.

Table 1
The characteristics 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

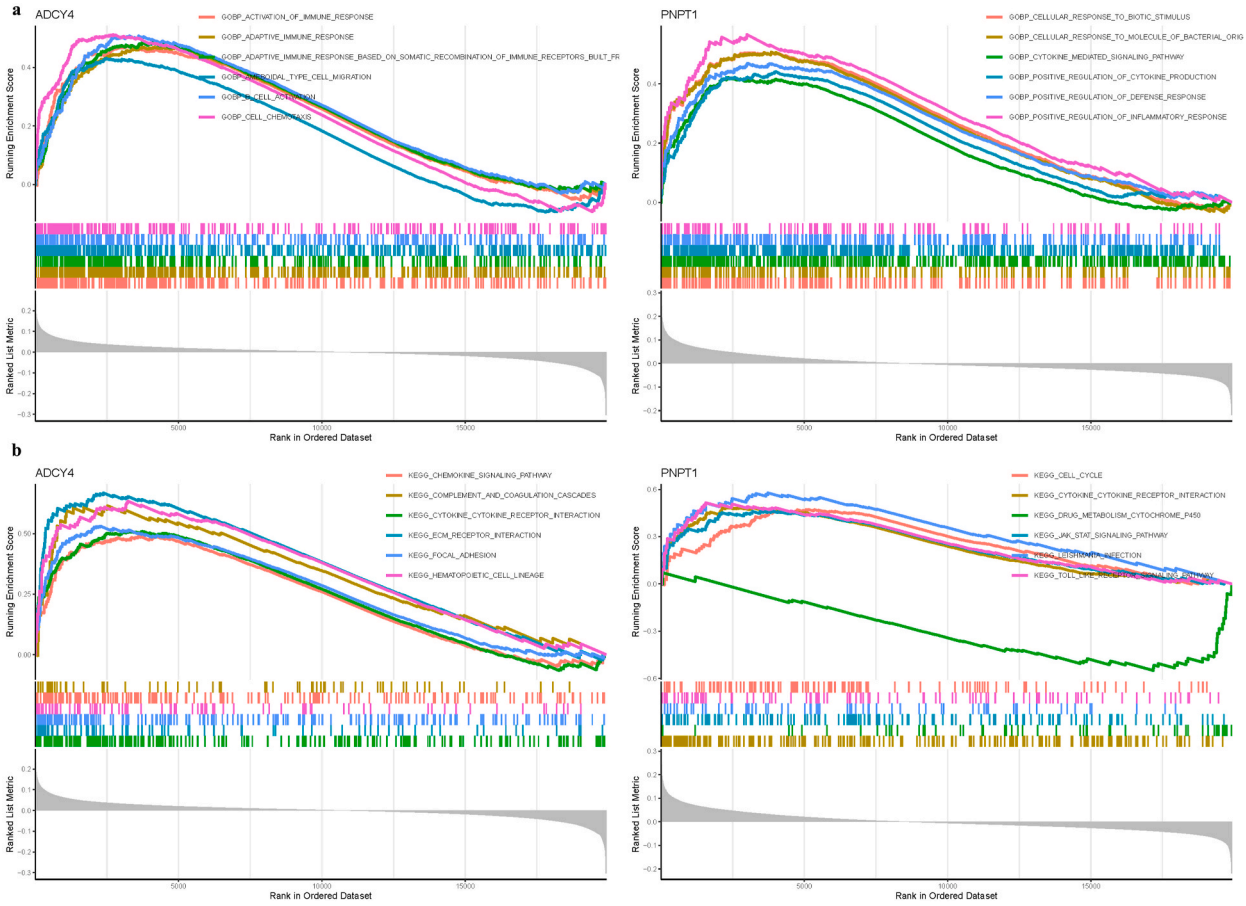


Fig. 5. GSEA of analysis in ADCY4 and PNPT1. (a): GO. (b): KEGG.

cascades, hematopoietic cell lineage, and the target genes of PNPT1 were involved in leishmania infection, toll-like receptor signaling pathway, and cytokine-cytokine receptor interaction (Fig. 5b and Table S5).

3.5. Analysis of immune cells

The next sub-section will start to unveil the intricate relationship of the immune microenvironment in the development of UC. In light of this, the immune cell dynamics were analyzed in UC in order to understand their role and interactions. To visualize these patterns a violplot was created and from this plot, it could be seen that the expression of the different immune cells was significantly different in control compared to treated animals. More importantly, the comparison of the control group and the experimental group through the violplot revealed the over-expression of MEs T cells CD4 memory resting, M2-Macrophages, Dendritic cells resting, and

Mast cells resting. On the other hand, the following were significantly up-regulated in the treatment group: Naïve B cells, Plasma cells, T cells CD4 Memory Activated, M0 Macrophages, M1 Macrophages, Dendritic cells activated, and Neutrophils. Such upregulation in the treatment group is very helpful for identifying the possible immunological activity or changes that might happen in the condition of UC. Moreover, to the given immunological landscape, a correlation analysis was added. This analysis was expected to demystify the relationship that the genes of interest have with various immune cells. Indeed, such views are important in an attempt to assess the part played by genetic factors in the formation or alteration of the immune milieu in UC. Based on these analyses the depicted Fig. 6a and b are suggestive of understanding the interaction of immune cells and genetic factors in UC.

3.6. GSVA

The GO analysis showed that ADCY4 was mainly involved in BP xenobiotic glucuronidation, BP brush border assembly, BP flavonoid glucuronidation, BP negative regulation of keratinocyte differentiation, MF lysophosphatidic acid phosphatase activity, MF arylesterase activity, BP ph elevation, BP cardiac muscle cell-cell adhesion, MF zinc-dependent alcohol dehydrogenase activity. PNPT1 was primarily enriched for the BP involved in the positive regulation of cell-cell adhesion by integrin, MF zinc-dependent alcohol dehydrogenase activity, MF arylesterase activity, BP ph elevation, BP uronic acid metabolic process, BP xenobiotic glucuronidation, MF lysophosphatidic acid phosphatase activity, CC microvesicle, and BP brush border assembly (Fig. 7a). ADCY4 gene/protein blocker was identified to modulate metabolism, including pyruvate metabolism, propanoate metabolism butanoate metabolism peroxisome and fatty acid metabolism valine, leucine, and isoleucine degradation and Citrate cycle TCA cycle and limonene and pinene degradation. The downregulated genes in PNPT1 were Pentose and glucuronate interconversions, Retinol metabolism, Alpha-linolenic Acid metabolism, Maturity-onset diabetes of the young, Metabolism of Xenobiotics by cytochrome p450, Drug metabolism Cytochrome

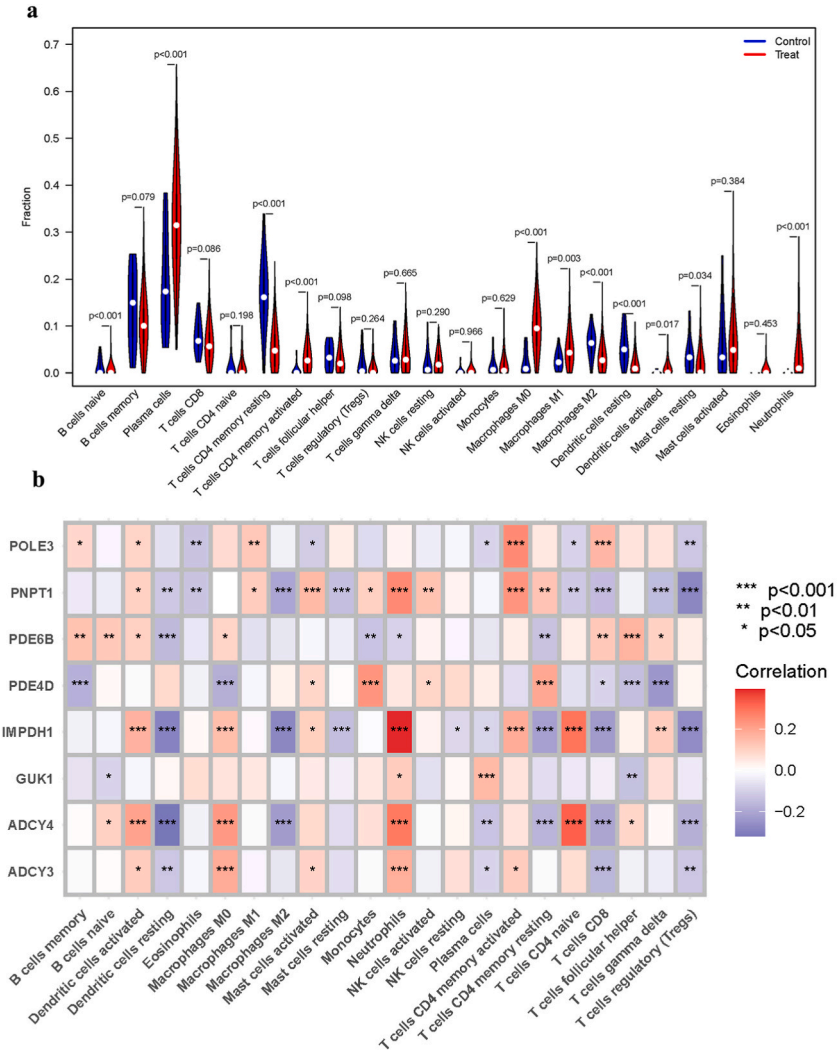


Fig. 6. Expression of immune cells. (a) Expression of immune cells in different clusters. (b) Correlation between PMGs and immune cells.

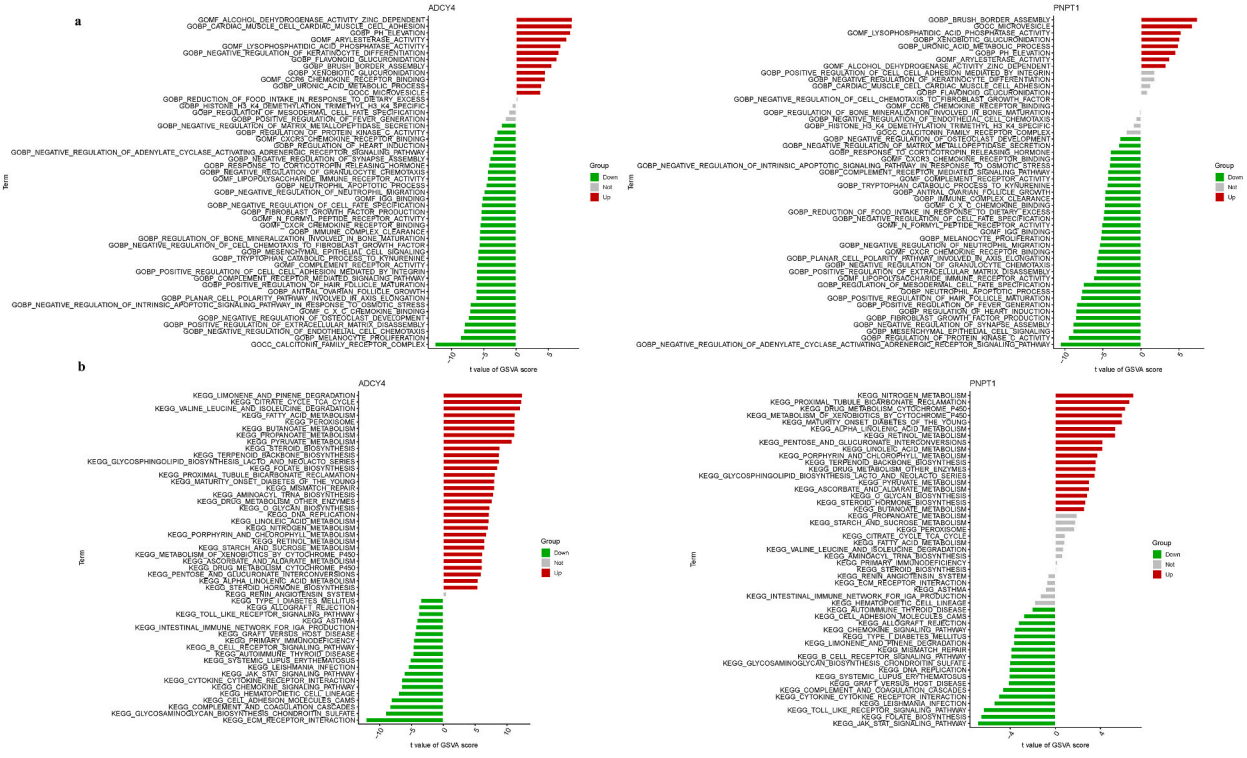


Fig. 7. GSVA of analysis in ADCY4 and PNPT1. (a): GO. (b): KEGG.

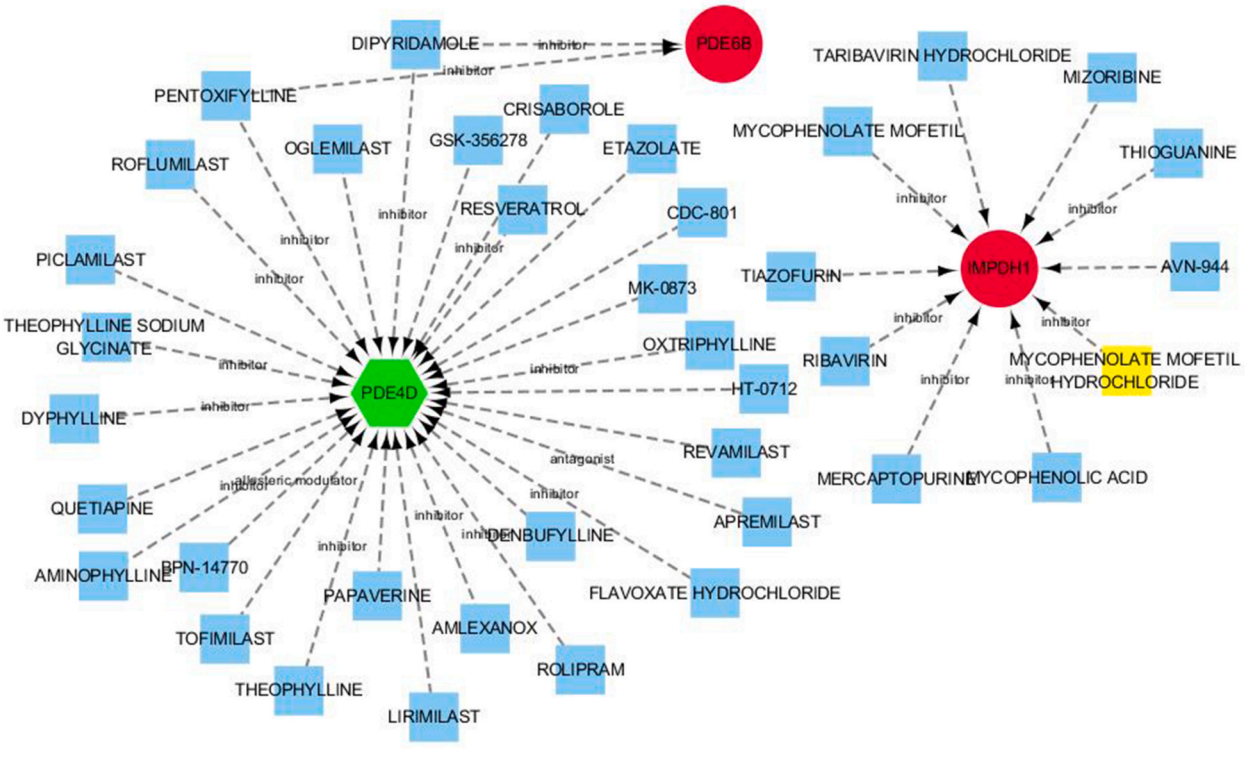


Fig. 8. Drug-gene interactions. Note: Filled red circles are up-regulation genes, green filled hexagonal are down regulation genes while blue boxes are linked drugs. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

p450, proximal Tubule Bicarbonate reclamation and Nitrogen metabolism (Fig. 7b).

3.7. Drug-gene interactions

The drugs were identified to have potential for interacting with the eight hub genes such as ribavirin, mercaptopurine, mycophenolate mofetil, mycophenolic acid, thioguanine, mycophenolate mofetil hydrochloride, tiazofurin, mizoribine, avn-944, taribavirin hydrochloride, dipyridamole (Table S6) (Fig. 8).

3.8. Identification of common RNAs and construction of miRNA-lncRNA shared genes network

Among all identified candidates from three databases, 232 miRNAs and 297 lncRNAs were implicated in UC (Tables S7a–b). Table S7 displays the target genes of these miRNAs in the corresponding miRNA databases. Some of the databases employed include miRanda [23] miRDB [24] and TargetScan [25]. If the corresponding database found the relevant miRNA, it gave a score of 1. From this, it will be seen that 3 points are awarded when all three databases can be matched. To ascertain the lncRNA data linked to the miRNA, the miRNA was cross-referenced with the spongeScan database [26]. This miRNA-lncRNA-gene network was built according to these mutual genes plus miRNA and lncRNA that were determined by Lasso regression and SVM-RFE. In the constructed network, the total of 206 lncRNA, 198 miRNA, and some common genes, which consisted of 7 central genes: POLE3, IMPDH1, PNPT1, PDE4D, ADCY3, PDE6B, and GUK1 (Fig. 9).

3.9. Validation of hub genes

To enhance the confidence and prediction accuracy of the model, the GSE179285 dataset was used for validation. The GSE179285 analysis further confirmed their potential relevance to UC (Fig. 10).

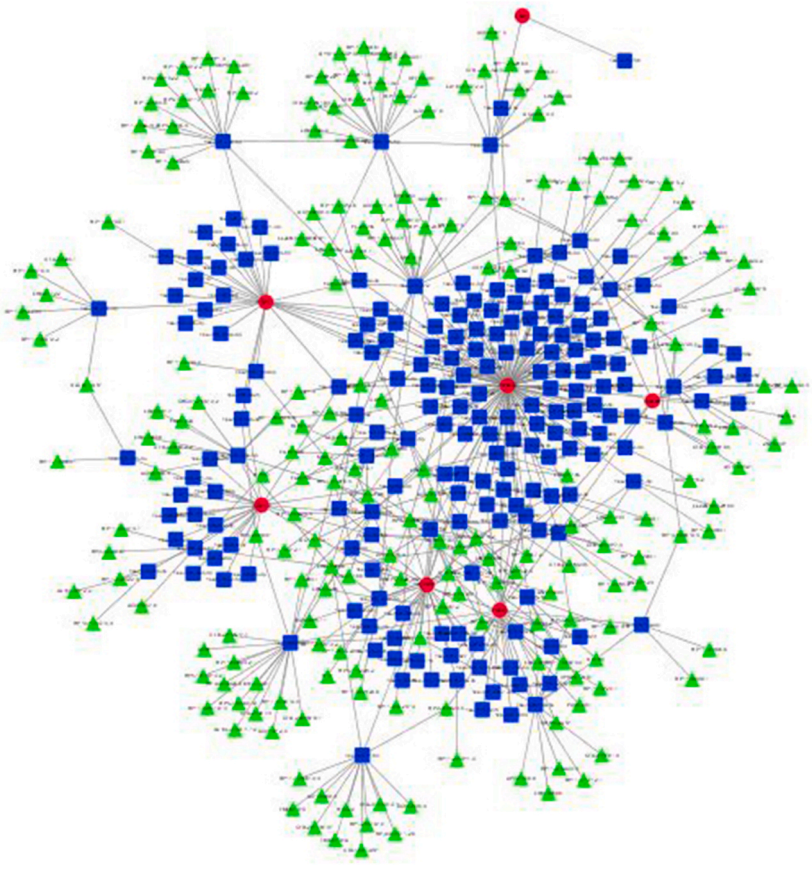


Fig. 9. miRNAs-LncRNAs shared genes network. Note: Red circles are mrnas, blue quadrangles are miRNAs, and green triangles are lncRNAs. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

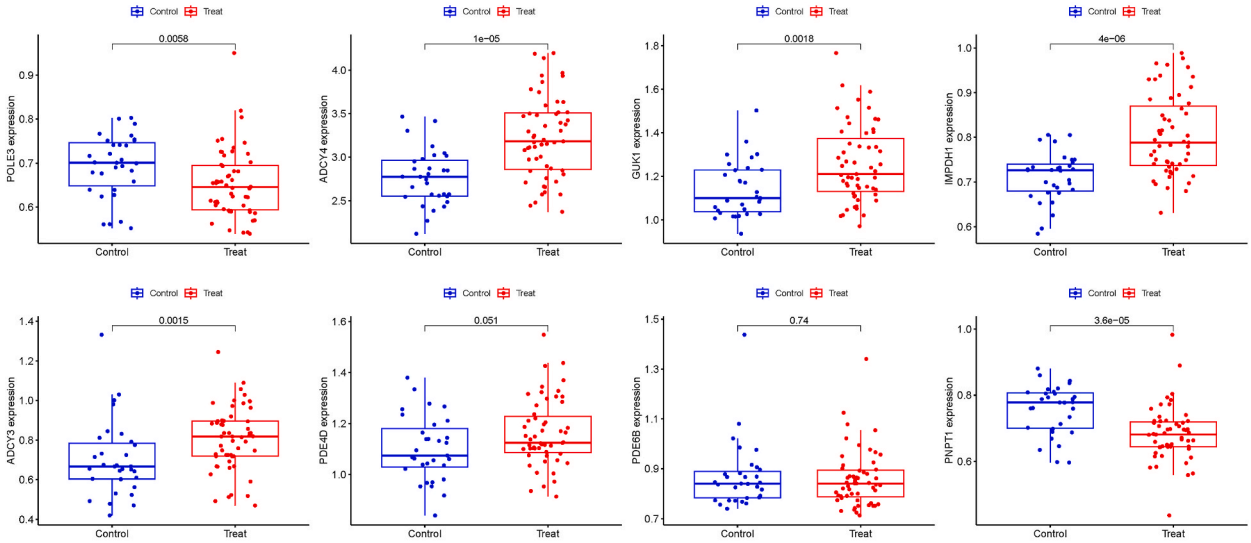


Fig. 10. Eight hub genes were validated.

3.10. Model verification

The Boxplots represented the residual expression level of these genes in UC as depicted in (Fig. 11a). It is possible to identify certain fluctuations in the ratios between the four different modes (Fig. 11b). The PMGs' diagnostic capacity in distinguishing UC from control samples revealed a satisfactory diagnostic value, with an AUC of RF: The PMGs' diagnostic capacity in distinguishing UC from control samples revealed a satisfactory diagnostic value, with an AUC of RF: Based on the given criteria the proposed research is as follows: 000; SVM: 0.999; XGB: 1.000; GLM: 1. For the first one the AUC was 0.000 (Fig. 11c), and for the second one, it was 0.000 (Fig. 11c). The HR was 931 (95 % CI 0.812–1.000) in GSE179285 (Fig. 11d).

3.11. Mendelian randomization analysis

Hence, in investigating the direct relationship between the ADCY4 and PNPT1 and UC incidence, a forest plot was adopted whereby the similarity of the data was observed from a general pattern of symmetry shown in the figure. Thus, using sensitivity analysis based on the "leave-one-out" approach, it was found that the IVW analysis was not significantly affected by the exclusion of individual SNP; this is evidence that other SNPs in the contribution reflected a similar result to the analysis based on the entire dataset. To further authenticate the outcomes, MR-Egger regression analysis was conducted, bolstering the integrity and reliability of our results and the chosen analytical framework (Fig. 10a and b). In addition, the Mendelian data of ADCY4 and PNPT1 were compared (Table 2).

3.12. Components-targets docking analysis

We performed molecular docking of ADCY4 and PNPT1 to the drug (Table 3 and Figs. 12 and 13).

4. Discussion

UC is actually a unique type of IBD whose etiology remains somewhat unknown, but it primarily involves the colorectum and common presentation includes bloody diarrhea along with abdominal pain [27]. Environmental factors that are involved include; genetic predisposition and dysfunction in gut microbiome [12]. UC is defined by mucosal inflammation starting at the rectum and going in a proximal manner in the colon without deep endoscopic involvement of the colon and limited to the mucosa, whereas Crohn disease disease affects the four layers of the bowel wall [28]. This inflammation is due to impaired intestinal physiologic and mucosal integrity and is usually characterized by a chronic and relapsing nature. UC patients may experience unpredictable relapses with fever, tenesmus, anemia, and in worse cases, intestinal perforation, thus making their quality of life very poor [29]. At present 5-aminosalicylic acid (5-ASA), rectal corticosteroids, and biological drugs are the key to UC treatment [30]. However, about 15 % of patients ultimately need restorative proctocolectomy with ileal pouch-anal anastomosis (IPAA) for Medically recalcitrant disease or colonic neoplasia [31]. As UC incidence and prevalence increase in the global arena, it is thus important to acknowledge the need to increase awareness of UC management and the costs as well. This therefore stresses the need to find therapeutic interventions for UC that have fewer effects on the body [32].

Early-stage diagnosis and proper identification of the condition known as ulcerative colitis or UC is important in this case. The regulation of gene expression especially concerning purine metabolism—a major component in the signal transduction processes, is

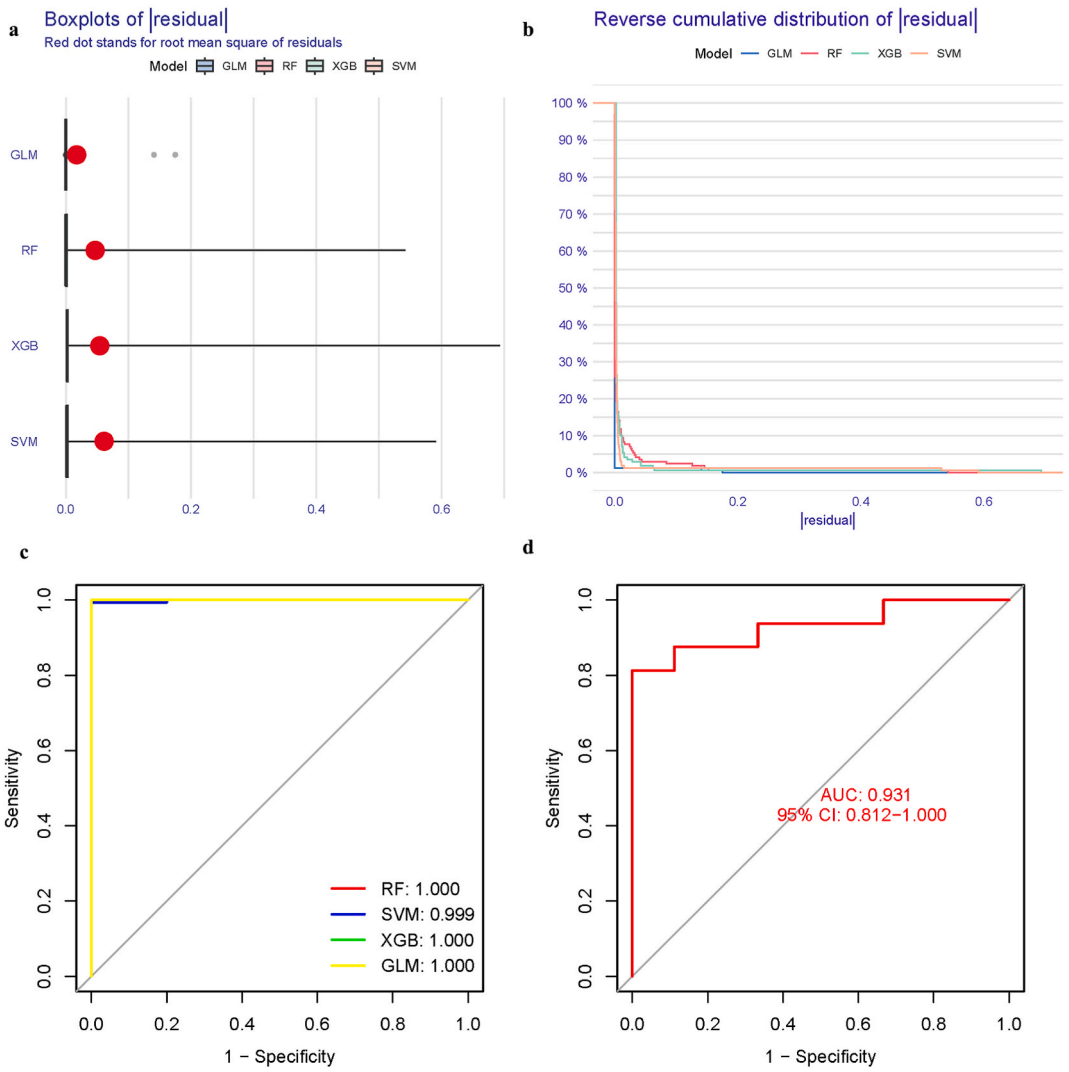


Fig. 11. Model verification. (a) Residual expression patterns. (b) Model expression patterns (c) AUC of model. (d) AUC of test group.

central. Abnormality of purine nucleotide metabolism may highly affect the switching on/off of a gene or a protein and also affect cell malignancy, invasiveness, and metastasis [33]. Though some risk markers have been found for several diseases, the values of these markers are still not very useful in clinical practice due to the lack of extensive reviews and huge sample validation. Until now, little research has been conducted investigating a set of purine metabolism regulators in cancer [34]. Nevertheless, the combined contribution of multiple purine metabolism-related genes in other pathologies, such as UC remains uninvestigated [35]. With progressing knowledge on the molecular basis of cancers, the emphasis is gradually being moved to non-cancerous components of diseases such as UC. New information might be obtained regarding the nature of UC by studying the various pathways of purine flare and their implication in the outcome. This might, in turn, reveal other targets that can be modulated for a substantial improvement in the management of UC.

Advanced bioinformatics approaches play a critical role in studying diseases like Ulcerative Colitis and colorectal cancer by enabling the integration and analysis of large-scale genetic, transcriptomic, and proteomic data. These approaches facilitate the identification of key biomarkers, such as the COL11A1 gene, which is implicated in the progression and prognosis of colorectal cancer. By understanding the molecular mechanisms and signaling pathways involved, bioinformatics can help in proposing targeted therapeutic interventions and improving patient-specific treatment strategies [36,37]. Through a comprehensive study, we successfully identified a group of 114 DEGs intricately associated with PMGs in UC. By implementing a robust methodological approach that seamlessly integrates DEG analysis, Lasso regression, and SVM-RFE, we were able to distinguish eight key PMGs: IMPDH1, GUK1, POLE3, ADCY3, ADCY4, PDE6B, PNPT1, and PDE4D. These hub genes were identified to have great discriminative values for UC and their importance in UC pathogenesis was further supported by validation with other datasets. However, our studies also showed that there is a lack of information on the specific transcription factors that regulate the expression of these genes in the context of the PMG.

Table 2
Mendelian randomization analysis.

Categories	ADCY4	PNPT1
id.exposure	prot-a-42	prot-a-42
id.outcome	ukb-b-19386	ukb-b-19386
outcome	Diagnoses - main ICD10: K51.9 Ulcerative colitis, unspecified || id:ukb-b-19386	Diagnoses - main ICD10: K51.9 Ulcerative colitis, unspecified || id:ukb-b-19386
exposure	Pituitary adenylate cyclase-activating polypeptide || id:prot-a-42	Pituitary adenylate cyclase-activating polypeptide || id:prot-a-42
method	Wald ratio	Wald ratio
nsnp	1	1
b	−0.00098612	−0.00098612
se	0.001004712	0.001004712
pval	0.326348507	0.326348507
lo_ci	−0.002955355	−0.002955355
up_ci	0.000983115	0.000983115
or	0.999014366	0.999014366
or_lci95	0.997049008	0.997049008
or_uci95	1.000983599	1.000983599

Table 3
The results of drug and disease molecular docking.

Protein	Ligand	Affinity energy/kcal/mol
ADCY4	TIAZOFURIN	−4.314
	TARIBAVIRIN HYDROCHLORIDE	−5.243
	BPN-14770	−4.136
	PICLAMILAST	−3.457
	DYPHYLLINE	−4.802
PNPT1	TIAZOFURIN	−3.578
	TARIBAVIRIN HYDROCHLORIDE	−4.892
	BPN-14770	−4.701
	PICLAMILAST	−3.221
	DYPHYLLINE	−4.384

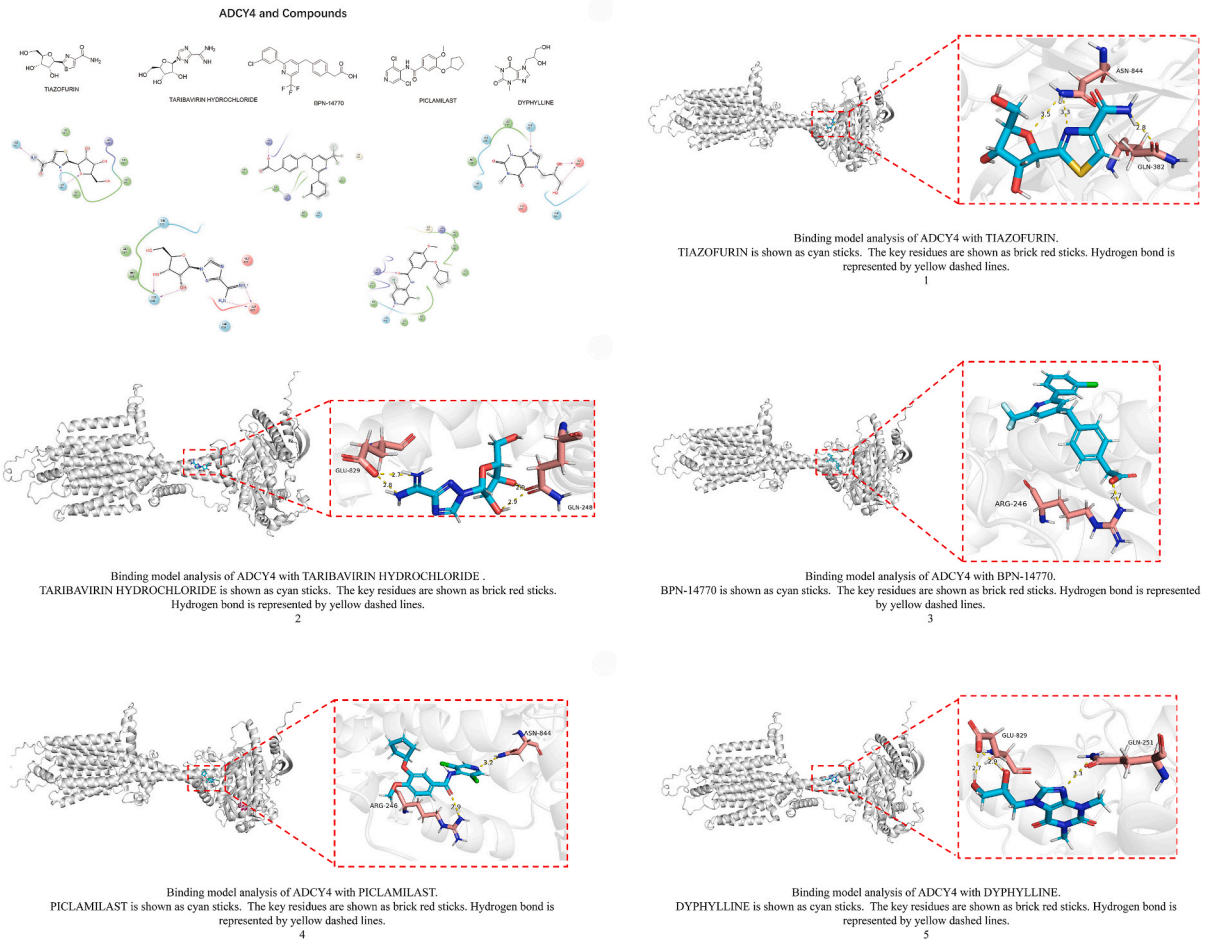


Fig. 12. The molecular docking model, active sites, and binding distances of ADCY4.

In the course of the current systematic review, focus was made on ADCY4 and PNPT1 which emerged as critical genes in the crosstalk between UC and PMGs. To elaborate on their biological functions, they were found to be involved in a variety of immune functions such as neutrophil-mediated immunity, activation of the immune response, and degranulation. From this analysis, a possible conclusion could be made that PMGs might have many regulatory impacts on multiple biological processes, with a major focus on immune-related processes. Hence, the regulatory impact of this discovery could potentially alter the natural course of the UC pathophysiological process. These genes may be critical in understanding the course of UC, which may lead to the identification of new targets for treatment and may provide a basis for new approaches to studying and treating this condition.

The study of purine metabolism, pivotal in maintaining cellular energy homeostasis and growth, has revealed profound implications in oncology and metabolic disorders. In the Fan research, a notable decrease in ADCY4 expression in breast cancer was observed, linked to promoter hypermethylation. Elevated ADCY4 levels were correlated with enhanced survival rates in breast cancer across

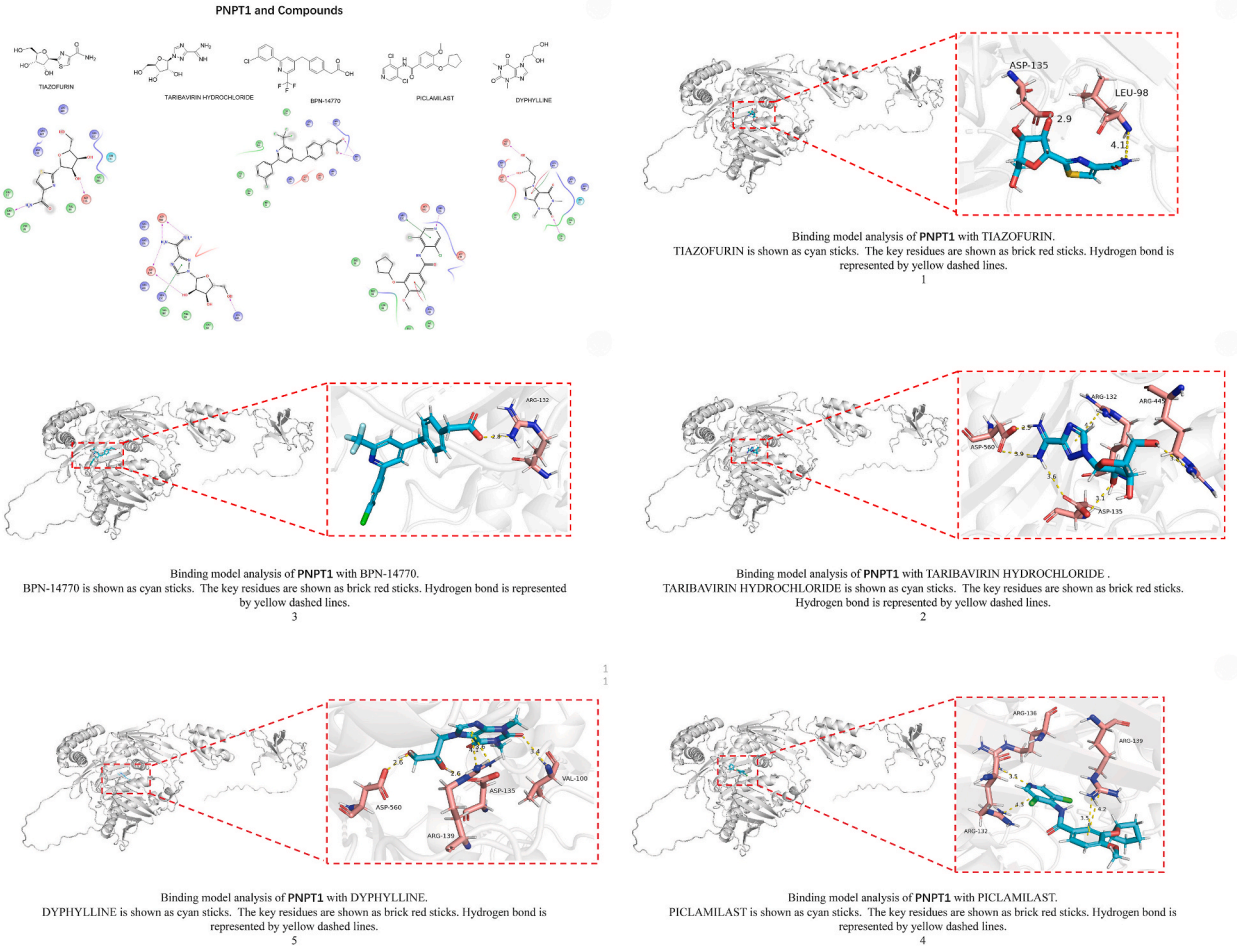


Fig. 13. The molecular docking model, active sites, and binding distances of PNPT1.

various intrinsic subtypes and tumor stages. This association is notably linked to G protein-coupled receptors and the downstream cAMP signaling pathway, with a marked enrichment in lysophosphatidic acid receptor 4 and glucagon-like peptide-1. ADCY4 emerges as a potential epigenetic biomarker and therapeutic target in breast cancer. The Liu study further underscores this by demonstrating that lower ADCY4 expression portends poorer survival in lung squamous cell carcinoma (LUSC), identifying it as a potential prognostic and therapeutic biomarker in LUSC. The study highlights a strong correlation between survival-related hub genes and LUSC tumorigenesis and progression. Similarly, Polyribonucleotide Nucleotidyltransferase 1 (Pnpt1) involvement in the maintenance of mitochondria is supported by its functions in mitochondrial RNA transportation, edition, and disposal. Mutations in Pnpt1 that have been associated with mitochondrial damage also initiate a type I interferon response pointing to its part in inflammation. Myeloid-specific Pnpt1-knockout mice showed increased IL-1β and IL-18 release in a sepsis model according to Hsu [38]. In addition, it was demonstrated that Pnpt1 controls NLRP3 inflammasome-mediated IL-1β secretion in macrophages which is a major finding to understand the regulation of inflammatory response. Also, the SCA25 phenotype with sensory and cerebellar ataxia has also been observed to be associated with a pathogenic variant in the gene PNPT1, which codes for Polyribonucleotide Nucleotidyltransferase PNPase 1 [39]. A splice variant from a large Australian family and a nonsense variant from an unrelated individual is found to produce different stop codons in the S1-domain section of PNPase. It is useful to note that this discovery adds to the more extended knowledge about the genetic causes of ataxias. Lastly, the data of the GSE179285 dataset clearly pointed out that some of the purine-related features can be used as prognostic factors in UC, which can be a future line in genomic studies. This is in addition to the growing realization of Nrf2 functions that go beyond the typical antioxidant response to oxidative stress. Such insights open the path to new approaches to UC treatment not only illustrating the potential of personalized approach in modern healthcare.

As a chronic manifestation of Inflammatory Bowel Disease (IBD), UC is distinctly marked by persistent inflammation of the colonic mucosa. Nonetheless, the etiology of this ailment, encompassing a complex network of unregulated immune responses and inflammatory mediators, remains partially understood [40]. Central to UC's pathophysiology is an aberrant immune reaction targeting the colonic mucosa. This involves the activation and infiltration of immune cells, notably T cells and B cells, into the colonic wall [41]. Specifically, a number of subgroups of helper T cells (Th), such as Th1, Th2, Th17 as well as Treg cells, play an important role in UC's

inflammatory process. Studies have shown that the levels of Th1 cytokines including interferon-γ and tumor necrosis factor-α are higher in UC [42]. At the same time, there are also changes in Th2 cytokines like an increase in the expression of interleukin-5 and interleukin 13. Subsequent investigations have also highlighted the role of Th17 cells, together with their related cytokines Il-17 and Il-22 in this situation [43]. Moreover, a decrease in the functionality of Treg cells means that they shall not be able to effectively stop an immune response which in turn increases inflammation. Besides the promotion of inflammation resulting from this immune activation, there is also enhanced production of inflammatory substances such as prostaglandins and leukotrienes causing further erosion of the mucosa and inflammation [44]. This inflammatory environment may be additionally enhanced by the injury to the mucosal barrier and by the pre-existing or developing dysregulation of immune responses to intestinal microbiota. Hence, we conclude that UC, in its complexity as a chronic immunopathological disorder, is characterized by immune cell activation, increased levels of inflammatory mediators, and altered immune-modulating systems [45]. In this regard, a vioplot visualization was used to define the expression patterns of the immune cells of interest. More importantly, a higher level of CD4 memory resting T cell, M2 macrophage, resting dendritic cell, and resting mast cell was found in the control group. On the other hand, Naïve B cells, plasma cells, activated CD4 memory T cells, M0 and M1 macrophages, activated dendritic cells, and neutrophils were up-regulated in the group that received treatment. These insights into UC's immune landscape underscore the imperative of deciphering immune pathways in devising novel therapeutic strategies. Immunomodulatory approaches, aiming to mitigate inflammatory processes and immune dysregulation in UC, represent a promising frontier in treatment methodologies. This highlights the criticality of immune pathways in both the comprehension and management of UC.

The burgeoning interest in delineating the nexus between UC and metabolic processes heralds a transformative era in contemporary medical research. The advent of advanced bioinformatics has catalyzed a paradigm shift, significantly enriching our comprehension of the molecular intricacies of UC and its associated pathologies [46–48]. This collective research endeavor plays a pivotal role in unraveling the molecular mechanisms underpinning UC and its diverse clinical manifestations. Focusing on the role of PMGs within the UC framework, our study addresses a significant lacuna in this domain. By harnessing comprehensive datasets from the GEO, particularly GSE206285, and GSE179285, our approach employed a suite of sophisticated analytical methodologies including GO, KEGG, and GSEA. These analytical techniques were instrumental in crafting an intricate predictive model that illuminates the complex involvement of PMGs in the pathogenesis of UC. Our research not only lays down a fundamental theoretical foundation but also paves the way for future inquiries into the realm of metabolic dysregulation in UC and the exploration of potential therapeutic strategies targeting these metabolic anomalies. Nevertheless, it is imperative to acknowledge that our study, despite its innovative stance, underscores the necessity for further empirical exploration to corroborate the primary mechanisms at play in UC. In spite of its theoretical and methodological advances, our study is not without limitations. While we employed comprehensive and refined bioinformatic techniques and validated our findings against other gene expression profiles, our analysis remains largely exploratory. In addition, the sample size, potential biases, and selective reporting are also possible limitations of this study. Further empirical research is necessary to substantiate these results so as to lay a theoretical foundation for subsequent investigations in this domain.

5. Conclusions

In the complex pathobiology of UC, the role of PMGs emerges as a critical element amidst a labyrinth of interactions involving various targets, pathways, signaling modalities, and regulatory mechanisms. PMGs, responsible for synthesizing vital molecules such as IMPDH1, GUK1, POLE3, ADCY3, ADCY4, PDE6B, PNPT1, and PDE4D, are integral to UC's molecular architecture. Notably, ADCY4 and PNPT1 stand out, significantly impacting the metabolic regulation and related processes in UC. These multifaceted interactions encapsulate the challenges and opportunities inherent in deepening our understanding and improving the treatment of UC.

Abbreviations

Abbreviations	Full name
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

Data availability statement

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

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Funding

This study was not supported by any funding.

CRediT authorship contribution statement

Tian Chen: Formal analysis. **Yiqiu Tao:** Formal analysis. **Qingyuan Wang:** Resources, Project administration. **Yanni Pei:** Formal analysis, Data curation. **Zhenhua Zhao:** Formal analysis, Data curation. **Wei Yang:** Formal analysis, Data curation. **Yafeng Lu:** Data curation.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e38403.

References

[1] C. Le Berre, S. Honap, L. Peyrin-Biroulet, Ulcerative colitis, Lancet (N. Am. Ed.) 402 (10401) (2023) 571–584.
[2] J.D. Feuerstein, A.C. Moss, F.A. Farraye, Ulcerative colitis, Mayo Clin. Proc. 94 (7) (2019) 1357–1373.
[3] R. Ungaro, S. Mehandru, P.B. Allen, L. Peyrin-Biroulet, J.F. Colombel, Ulcerative colitis, Lancet (N. Am. Ed.) 389 (10080) (2017) 1756–1770.
[4] R. Patra, A.K. Dey, S. Mukherjee, Identification of genes critical for inducing ulcerative colitis and exploring their tumorigenic potential in human colorectal carcinoma, PLoS One 18 (8) (2023) e0289064, https://doi.org/10.1371/journal.pone.0289064. Published 2023 Aug 3.
[5] J.P. Segal, J.F. LeBlanc, A.L. Hart, Ulcerative colitis: an update, Clin. Med. 21 (2) (2021) 135–139.
[6] E.F. Stange, Mitochondria in ulcerative colitis, Cell Mol Gastroenterol Hepatol 12 (1) (2021) 352–353.
[7] M.F. Neurath, M. Leppkes, Resolution of ulcerative colitis, Semin. Immunopathol. 41 (6) (2019) 747–756.
[8] V. Rao, E. Cockburn, J.P. Segal, Is ulcerative colitis really ever "cleared?", Inflamm. Bowel Dis. (2023).
[9] Y. Zhu, S. Yang, N. Zhao, C. Liu, F. Zhang, Y. Guo, H. Liu, CXCL8 chemokine in ulcerative colitis, Biomed. Pharmacother. 138 (2021) 111427.
[10] D.H. Jo, J.H. Kim, J.H. Kim, Tumor environment of retinoblastoma, intraocular cancer, Adv. Exp. Med. Biol. 1296 (2020) 349–358, https://doi.org/10.1007/978-3-030-59038-3_21 [Journal Article].
[11] L. Xia, L. Oyang, J. Lin, S. Tan, Y. Han, N. Wu, et al., The cancer metabolic reprogramming and immune response. [Journal Article; Research Support, Non-U.S. Gov't; Review], Mol. Cancer 20 (1) (2021) 28, https://doi.org/10.1186/s12943-021-01316-8.
[12] Y. Zhang, Z. Zhang, The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. [Historical Article; Journal Article; Research Support, Non-U.S. Gov't; Review], Cell. Mol. Immunol. 17 (8) (2020) 807–821, https://doi.org/10.1038/s41423-020-0488-6.
[13] S. Yang, B. Zhang, W. Tan, L. Qi, X. Ma, X. Wang, A novel purine and uric metabolism signature predicting the prognosis of hepatocellular carcinoma, Front. Genet. 13 (2022) 942267.
[14] J. Liu, S. Hong, J. Yang, et al., Targeting purine metabolism in ovarian cancer, J. Ovarian Res. 15 (1) (2022) 93.
[15] J. Yin, W. Ren, X. Huang, J. Deng, T. Li, Y. Yin, Potential mechanisms connecting purine metabolism and cancer therapy, Front. Immunol. 9 (2018) 1697.
[16] O.P. Shatova, E.V. Butenko, E.V. Khomutov, D.S. Kaplun, I.E. Sedakov, I.I. Zinkovych, Metformin impact on purine metabolism in breast cancer, Biomed. Khim. 62 (3) (2016) 302–305.
[17] X. Chen, J. Chen, miR-10b-5p-mediated upregulation of PIEZO1 predicts poor prognosis and links to purine metabolism in breast cancer, Genomics 114 (3) (2022) 110351.
[18] Y. Yuan, N. Li, M. Fu, M. Ye, Identification of critical modules and biomarkers of ulcerative colitis by using WGCNA, J. Inflamm. Res. 16 (2023) 1611–1628.
[19] M. Xu, Y. Kong, N. Chen, W. Peng, R. Zi, M. Jiang, J. Zhu, Y. Wang, J. Yue, J. Lv, et al., Identification of immune-related gene signature and prediction of CeRNA network in active ulcerative colitis, Front. Immunol. 13 (2022) 855645.
[20] Y. Wu, X. Liu, G. Li, Integrated bioinformatics and network pharmacology to identify the therapeutic target and molecular mechanisms of Huangqin decoction on ulcerative Colitis, Sci. Rep. 12 (1) (2022) 159.
[21] Z. Wu, Y. Gao, L. Cao, Q. Peng, X. Yao, Purine metabolism-related genes and immunization in thyroid eye disease were validated using bioinformatics and machine learning, Sci. Rep. 13 (1) (2023 Oct 26) 18391.
[22] Z. Wu, N. Li, Y. Gao, L. Cao, X. Yao, Q. Peng, Glutamine metabolism-related genes and immunotherapy in nonspecific orbital inflammation were validated using bioinformatics and machine learning, BMC Genom. 25 (1) (2024 Jan 17) 71.
[23] T.R. De Carvalho, A.A. Giaretta, B.F. Teixeira, L.B. Martins, New bioacoustic and distributional data on Bokermannohyla sapiranga Brandao et al., 2012 (Anura: Hylidae): revisiting its diagnosis in comparison with B. pseudopseudis (Miranda-Ribeiro, 1937), Zootaxa 3746 (2013) 383–392.
[24] Y. Chen, X. Wang, miRDB: an online database for prediction of functional microRNA targets, Nucleic Acids Res. 48 (D1) (2020) D127–D131.
[25] D. Mon-Lopez, C.M. Tejero-Gonzalez, Validity and reliability of the TargetScan ISSF Pistol & Rifle application for measuring shooting performance, Scand. J. Med. Sci. Sports 29 (11) (2019) 1707–1712.

[26] P. Furio-Tari, S. Tarazona, T. Gabaldon, A.J. Enright, A. Conesa, spongeScan: a web for detecting microRNA binding elements in lncRNA sequences, Nucleic Acids Res. 44 (W1) (2016) W176–W180.
[27] S. Ghoneim, Etrasimod in ulcerative colitis, Gastroenterology 165 (5) (2023) 1302–1303.
[28] F.S. Macaluso, Orlando A: ozanimod for ulcerative colitis, N. Engl. J. Med. 386 (2) (2022) 194.
[29] A.N. Ananthakrishnan, Upadacitinib for ulcerative colitis, Lancet (N. Am. Ed.) 399 (10341) (2022) 2077–2078.
[30] J. Pravda, Can ulcerative colitis be cured? Discov. Med. 27 (149) (2019) 197–200.
[31] G. Qi, Y. Mi, X. Shi, H. Gu, R.D. Brinton, F. Yin, ApoE4 impairs neuron-astrocyte coupling of fatty acid metabolism, Cell Rep. 34 (1) (2021) 108572.
[32] K. Yang, J. Li, L. Tao, Purine metabolism in the development of osteoporosis, Biomed. Pharmacother. 155 (2022) 113784, https://doi.org/10.1016/j.biopha.2022.113784.
[33] M. Furuhashi, New insights into purine metabolism in metabolic diseases: role of xanthine oxidoreductase activity. [Journal Article; Research Support, Non-U.S. Gov't; Review], Am. J. Physiol. Endocrinol. Metab. 319 (5) (2020) E827–E834, https://doi.org/10.1152/ajpendo.00378.2020.
[34] F. Yu, F. Quan, J. Xu, Y. Zhang, Y. Xie, J. Zhang, et al., Breast cancer prognosis signature: linking risk stratification to disease subtypes. [Journal Article; Research Support, Non-U.S. Gov't; Review], Briefings Bioinf. 20 (6) (2019) 2130–2140, https://doi.org/10.1093/bib/bby073.
[35] Z. Jiang, H. Shen, B. Tang, Q. Yu, X. Ji, L. Wang, Quantitative proteomic analysis reveals that proteins required for fatty acid metabolism may serve as diagnostic markers for gastric cancer, Clin. Chim. Acta 464 (2017) 148–154.
[36] R. Patra, J. Chakraborty, N.C. Das, S. Mukherjee, An integrated omics study on the role of HDAC9 gene in the oncogenic events of human gastrointestinal-tract associated cancers, Hum. Genet. 37 (2023) 201189, https://doi.org/10.1016/j.humgen.2023.201189.
[37] R. Patra, N.C. Das, S. Mukherjee, Exploring the differential expression and prognostic significance of the COL11A1 gene in human colorectal carcinoma: an integrated bioinformatics approach, Front. Genet. 12 (2021) 608313, https://doi.org/10.3389/fgene.2021.608313.
[38] M. Barbier, M. Bahlo, A. Pennisi, M. Jacoupy, R.M. Tankard, C. Ewenczyk, K.C. Davies, P. Lino-Coulon, C. Colace, H. Rafehi, et al., Heterozygous PNPT1 variants cause spinocerebellar ataxia type 25, Ann. Neurol. 92 (1) (2022) 122–137.
[39] C.G. Hsu, W. Li, M. Sowden, C.L. Chavez, B.C. Berk, Pnpt1 mediates NLRP3 inflammasome activation by MAVS and metabolic reprogramming in macrophages, Cell. Mol. Immunol. 20 (2) (2023) 131–142.
[40] D. Noviello, R. Mager, G. Roda, R.G. Borroni, G. Fiorino, S. Vetrano, The IL23-IL17 immune Axis in the treatment of ulcerative colitis: successes, defeats, and ongoing challenges, Front. Immunol. 12 (2021) 611256.
[41] Z. Pan, H. Lin, Y. Fu, F. Zeng, F. Gu, G. Niu, J. Fang, B. Gu, Identification of gene signatures associated with ulcerative colitis and the association with immune infiltrates in colon cancer, Front. Immunol. 14 (2023) 1086898.
[42] J. Huang, J. Zhang, F. Wang, B. Zhang, X. Tang, Comprehensive analysis of cuproptosis-related genes in immune infiltration and diagnosis in ulcerative colitis, Front. Immunol. 13 (2022) 1008146.
[43] V. Mitsialis, S. Wall, P. Liu, J. Ordovas-Montanes, T. Parmet, M. Vukovic, D. Spencer, M. Field, C. McCourt, J. Toothaker, et al., Single-cell analyses of colon and blood reveal distinct immune cell signatures of ulcerative colitis and Crohn's disease, Gastroenterology 159 (2) (2020) 591–608.
[44] B.S. Boland, Z. He, M.S. Tsai, J.G. Olvera, K.D. Omilusik, H.G. Duong, E.S. Kim, A.E. Limary, W. Jin, J.J. Milner, et al., Heterogeneity and clonal relationships of adaptive immune cells in ulcerative colitis revealed by single-cell analyses, Sci Immunol 5 (50) (2020).
[45] M. Xu, Y. Kong, N. Chen, W. Peng, R. Zi, M. Jiang, J. Zhu, Y. Wang, J. Yue, J. Lv, et al., Identification of immune-related gene signature and prediction of CeRNA network in active ulcerative colitis, Front. Immunol. 13 (2022) 855645.
[46] Z. Pan, H. Lin, Y. Fu, F. Zeng, F. Gu, G. Niu, J. Fang, B. Gu, Identification of gene signatures associated with ulcerative colitis and the association with immune infiltrates in colon cancer, Front. Immunol. 14 (2023) 1086898.
[47] W. Wang, X. Gao, N. Kang, C. Wang, C. Li, H. Yu, X. Zhang, Shared biomarkers and immune cell infiltration signatures in ulcerative colitis and nonalcoholic steatohepatitis, Sci. Rep. 13 (1) (2023) 18497.
[48] Z. Wu, Z. Cai, H. Shi, X. Huang, M. Cai, K. Yuan, P. Huang, G. Shi, T. Yan, Z. Li, 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) 14 (8) (2022) 3569–3596.