


Research

Prostate cancer and metabolic syndrome: exploring shared signature genes through integrative analysis of bioinformatics and clinical data

Maomao Guo¹  · Sudong Liang¹ · Zhenghui Guan¹ · Jingcheng Mao² · Zhibin Xu¹ · Wenchao Zhao¹ · Hao Bian¹ · Jianfeng Zhu²  · Jiangping Wang¹ · Xin Jin¹ · Yuan Xia^{2,3} 

Received: 14 May 2024 / Accepted: 5 May 2025

Published online: 08 May 2025

© The Author(s) 2025 

Abstract

The incidence of both prostate cancer (PCa) and metabolic syndrome (MS) has been steadily increasing due to changes in population structure and lifestyle. These two conditions frequently co-occur, yet their shared pathogenic mechanisms remain unclear. In this study, we utilized bioinformatics and machine learning techniques to analyze public datasets and validated our findings using clinical specimens from our center to identify common signature genes between PCa and MS. We began by screening differentially expressed genes (DEGs) and module genes through Linear models for microarray analysis (Limma) and Weighted Gene Co-expression Network Analysis (WGCNA) of four microarray datasets from the GEO database (PCa: GSE8511, GSE32571, and GSE104749; MS: GSE98895). Comprehensively bioinformatics analyses, including functional enrichment, LASSO, and random forest algorithms, coupled with receiver operating characteristic (ROC) and precision recall curve (PRC) analyses were conducted. We identified 423 DEGs in the PCa dataset and 2481 differentially modular genes in the MS dataset. Among these, 52 intersection genes enriched in immunomodulatory pathways were found. Three common signature genes, namely GPD1L, ACY1, and C12orf75, were identified through LASSO and random forest analyses. Subsequent validation using clinical specimens confirmed differential expression of these genes in PCa, with survival analysis indicating that elevated expression of ACY1 is associated with adverse prognosis in PCa patients. Additionally, immunoinfiltration analysis revealed higher levels of macrophage M0 and activated dendritic cells in PCa tissues. In summary, our study identifies three shared signature genes between PCa and MS, with ACY1 demonstrating adverse prognostic significance in PCa. Our findings provide a foundation for elucidating the pathogenic mechanisms and interplay between PCa and MS, offering novel insights for identifying potential therapeutic targets in PCa.

Keywords Prostate cancer · Metabolic syndrome · Bioinformatics analysis · Immune infiltration

Maomao Guo, Sudong Liang and Zhenghui Guan contributed equally to this work.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12672-025-02561-9>.

✉ Xin Jin, jxin9495@126.com; ✉ Yuan Xia, xia816yuan@163.com | ¹Department of Urology, The Affiliated Taizhou People's Hospital of Nanjing Medical University, Taizhou 225300, China. ²Department of Hematology, The Affiliated Taizhou People's Hospital of Nanjing Medical University, Taizhou 225300, China. ³Department of Hematology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China.



1 Introduction

Prostate cancer (PCa) ranks as the second most common malignancy among males worldwide, according to the 2020 Global Cancer Statistics [1]. Data from the United States in 2023 show that PCa has the highest incidence (29%) and the second-highest mortality (11%) among male malignancies. Notably, the annual incidence of PCa exhibited a 3% increase from 2014 to 2019, leading to an additional 99,000 new cases [2]. Changes in population structure (aging population) and lifestyle modifications (high-energy diet and reduced physical activity) are contributing factors to the increasing incidence of PCa. These changes also result in metabolic imbalances, leading to the rising prevalence of metabolic syndrome (MS) over the past few decades [3].

MS is a cluster of metabolic disorders characterized by disturbances in the metabolism of proteins, fats, and carbohydrates, manifested by features such as obesity, hyperglycemia, dyslipidemia, and hypertension [4]. Previous studies have shown a correlation between MS and the risk of developing high-grade disease, and adverse pathology in PCa [5–7]. Adipose tissue, functioning as an endocrine organ, modulates the synthesis and bioavailability of various hormones. Hormones such as androgens and adiponectin, along with overweight, are closely associated with the occurrence and progression of PCa [3].

Although the mechanisms of PCa remain incompletely understood, considerable research indicates that its etiology may involve alterations in the immune microenvironment, chronic inflammation, oxidative stress, and specific molecular aberrations in molecules like the androgen receptor and TP53 [8]. Importantly, alterations in the immune microenvironment and inflammation play crucial roles in the development of MS [9]. Thus, a specific immune microenvironment may be a pivotal factor in the pathogenesis of both MS and PCa, influencing the progression and prognosis of PCa. However, current research on the association and underlying mechanisms between MS and PCa remains limited. Therefore, elucidating the genetic characteristics of MS accompanied by PCa (PCa-MS) and identifying potential biomarkers holds significant clinical relevance.

In this study, we utilized advanced bioinformatics and machine learning techniques to analyze the genetic characteristics of PCa and MS, aiming to identify signature genes and underlying mechanisms of PCa-MS. In an era where the incidence of both PCa and MS is gradually rising, especially given their mutual association and frequent co-occurrence, this study provides a new perspective for exploring potential targets and the immune landscape in PCa and MS.

2 Material and methods

2.1 Patients

PCa tumor samples and non-neoplastic prostate tissue were collected from patients diagnosed with PCa and benign prostatic hyperplasia (BPH) at the Department of Urology, Affiliated Taizhou People's Hospital of Nanjing Medical University, from January 2023 to March 2023. Diagnoses of PCa and BPH were made according to the World Health Organization (WHO) guidelines. This study was conducted in accordance with the Declaration of Helsinki and was approved by the institutional review boards of the Affiliated Taizhou People's Hospital of Nanjing Medical University Ethics Committee (No. KY2023-011–01). Informed consent was obtained from patients for tissue use.

2.2 Raw data acquisition and data processing

Three microarray datasets (GSE8511, GSE32571, and GSE104749) of PCa and one microarray dataset (GSE98895) of MS, including expression profile matrix and clinical information, were downloaded from the NCBI Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>). Datasets includes a minimum of 8 samples, encompassing both human normal and abnormal tissues, were chosen for subsequent analysis. All details pertaining to the selected datasets are consolidated in Supplementary Table 1. R software (<https://www.r-project.org/>) was used to pre-process the datasets. Firstly, the datasets were background corrected, normalized, and log2 transformed using the "Affy" package (<https://www.bioconductor.org/packages/release/bioc/html/affy.html>). The gene expression

corresponding to the multiple probes was summarized. Subsequently, the three PCa datasets were merged using the R package "inSilicoMerging". Methods from Johnson WE et al. [10] were adopted to remove the batch effects. The Schematic overview of this study is presented in Fig. 1A.

2.3 Analysis of differentially expressed genes (DEGs)

Differential analysis was performed using the R package "Limma" (v3.40.6) [11] to obtain DEGs between cancer and control groups. Specifically, the processed datasets of PCa and MS were subjected to multiple linear regression using the "LMFit" function. Moderated t-statistics, moderated F-statistic, and log-odds of differential expression were computed through empirical Bayes moderation of the standard errors towards a typical value using the "eBays" function to determine the differential significance of each gene. A threshold for screening DEGs was set at an adjusted P-value < 0.05 and an absolute fold-change value > 1.5. Heatmaps and volcano plots were generated using the "heatmap" and "ggplot2" packages in R software, respectively.

2.4 Weighted gene co-expression network analysis (WGCNA)

WGCNA is a systems biology approach aimed at characterizing gene association patterns between samples. It is used to identify highly synergistic sets of genes and pinpoint candidate biomarker genes or therapeutic targets based on the endogeneity of the gene set and its association with the phenotype. Firstly, the Median Absolute Deviation (MAD) of each gene was calculated individually using the gene expression profile. The top 50% of genes with the smallest MAD were excluded, followed by the removal of outlier genes and samples using the "goodSamplesGenes" method of the R package "WGCNA". Subsequently, a scale-free co-expression network was constructed using WGCNA. Referring to

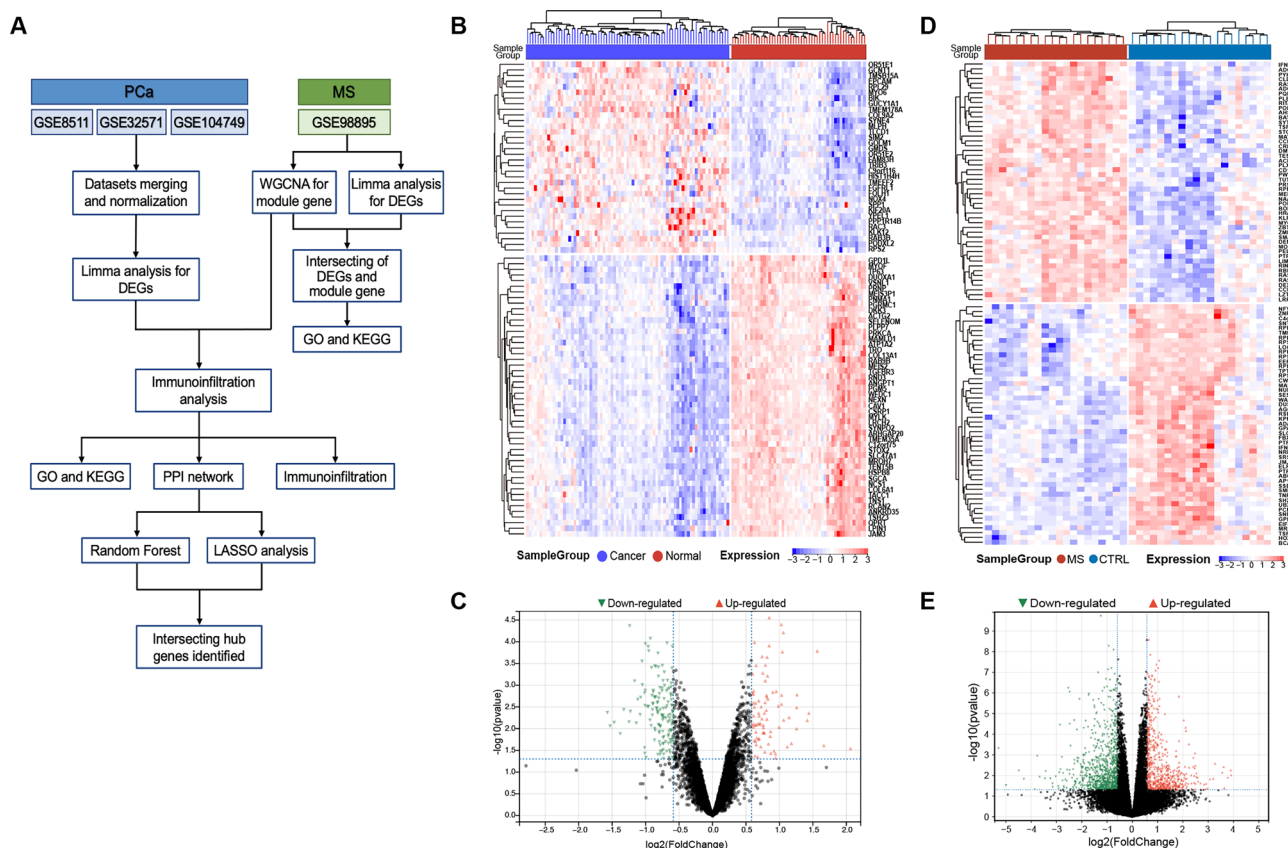


Fig. 1 Schematic overview of this study and the identification of DEGs. **A** Flowchart of the experimental design of this study. Heatmap (**B**) and volcano plots (**C**) representing DEGs between PCa and BPH patients in the merged datasets including GSE8511, GSE32571, and GSE104749. Heatmap (**D**) and volcano plots (**E**) of DEGs between MS patients and healthy subjects in the MS dataset (GSE98895)

previous studies [12], we ultimately identified eight co-expression modules; notably, the grey module was considered a set of genes that could not be assigned to any module.

2.5 Protein–protein interaction (PPI) analysis

A PPI network was constructed for 52 DEGs identified in PCa, utilizing data from the STRING database and setting a median confidence score threshold of 0.4 [13]. In this network, each node represents a unique protein, molecule, or gene, with the connecting edges denoting interactions between these biomolecular entities. Essential interacting genes were identified through the utilization of the MCODE plugin in Cytoscape, followed by their arrangement based on node degrees.

2.6 GO and KEGG analysis

The enrichment analysis was performed using the R package "clusterProfiler" (v3.14.3) [14] and visualized through the Sangerbox platform (<http://vip.sangerbox.com/>). The KEGG API (<https://www.kegg.jp/kegg/rest/keggapi.html>) was used to obtain canonical KEGG pathway gene notes, and the R package "org.Hs.eg.db" (v3.1.0) was used to obtain GO terms. Significance thresholds were set at a minimum gene set size of five and a maximum gene set size of 5000, with statistical significance defined by a *P* value < 0.05 and a false discovery rate (FDR) < 0.25.

2.7 LASSO analysis and random forest (RF) analysis

LASSO is a linear regression method using L1 regularization. If L1 regularization is selected, the weights of some learned features will be zero for sparse and feature selection purposes [15]. In this study, the R package "glmnet" [16] was used to consolidate survival time, survival state, and gene expression data, and regression analysis was conducted using the lasso-cox methodology. Moreover, a tenfold cross-validation was applied to acquire the optimum model. The Lambda value was set to 0.0249269053653555, resulting in 21 genes; RF is an integrated algorithm composed of decision trees belonging to the Bagging type. Integrating multiple weak classifiers was averaged, resulting in the overall model having excellent accuracy, generalization performance, and solid stability. The R package "randomForest" was used to perform RF analysis [17], and the crossover genes of LASSO and RF were regarded as promising targets for PCa diagnosis.

2.8 Receiver operating characteristic (ROC) evaluation

The R package "pROC" was used for ROC analysis to assess the utility of potential genes in the diagnosis of PCa (v1.17.0.1). Specifically, we obtained LASSO-RF intersection gene data, performed ROC analysis using the roc function of pROC, and used the ci function of pROC to assess the AUC and confidence intervals to get the AUC results. AUC > 0.7 was regarded as the optimal diagnostic result.

2.9 Immunohistochemical (IHC) staining

Tissue paraffin slices were deparaffinized in xylene and then rehydrated in a graded series of ethanol. The slices were then immersed in 3% H₂O₂ to inhibit the activity of endogenous peroxidase. To mitigate nonspecific binding, 10% normal rabbit serum was applied and allowed to incubate at room temperature for 30 min. After antigen retrieval and blocking, the slides were incubated with the designated antibodies overnight at 4 °C, followed by incubation with corresponding secondary antibodies for 1 h at room temperature, and subsequent treatment with the streptavidin peroxidase complex. The tissue sections were then stained using a 3,3-diaminobenzidine (DAB) substrate kit, with hematoxylin employed as the counterstain. Finally, images were collected using a light microscope.

2.10 Determining the optimal cut-off value for survival analysis

The optimal cut-off values for ACY1, C12orf75, and GPD1L expression levels were determined using the "surv_cutpoint" function from the R package "survminer". This function uses maximally selected rank statistics to calculate the most optimal cut-off for continuous variables based on log-rank statistics.

2.11 Immune cell infiltration analysis

CIBERSORT, a commonly used method for calculating immune cell infiltration, uses linear support vector regression-based deconvolution of the expression matrix of human immune cell subtypes. The abundance of the 22 immune cell species was assessed using the CIBERSORT package in R. Bar charts and heat maps represent the relative abundance of 22 immune cells in each sample. The relationships between immune cells were analyzed and displayed using correlation heat maps.

2.12 Data statistical analysis

Mann–Whitney U or Kruskal–Wallis tests were used for comparing continuous variables, and χ^2 or Fisher's exact tests were used for comparing categorical variables. Progression-free survival (PFS) and overall survival (OS) were plotted as Kaplan–Meier curves and compared by log-rank test. Data were analyzed using SPSS (v23.0) and R software (v4.1.1). All tests were two-tailed, and $P < 0.05$ was considered statistically significant.

3 Results

3.1 Data acquisition and pre-processing

Three datasets (GSE8511, GSE32571, and GSE104749) from the GEO database were merged into one integrated dataset to remove batch effects according to previous study [10, 18, 19]. In the realm of disease genomics, intersecting genes may elucidate the mechanisms underlying disease onset and progression. In this study, we have analyzed the co-expression profiles across diverse disease samples, aiming to further unveil the shared molecular characteristics among different diseases and uncover potential therapeutic targets. The initial upset plot showed 6000 intersection genes among the three PCa datasets (Supplementary Fig. 1 A). Subsequent boxplots, density maps, and UMAP unveiled noticeable differences in sample distribution among the datasets. After batch effect removal, the data distribution tended to be consistent among the datasets, indicating successful mitigation of the batch effect (Supplementary Fig. 1B–D).

3.2 Screening of differentially expressed genes

Linear models for microarray analysis (Limma) were conducted on both the PCa and MS datasets to identify DEGs. The heatmap and volcano plot representations illustrated that, in the merged PCa dataset, 423 genes were identified as differentially expressed between PCa and BPH patients, with 125 up-regulated and 298 down-regulated genes showing at least a 1.5-fold change ($FDR < 0.05$) (Fig. 1B, C). Analysis of the MS datasets revealed 1467 genes were identified as differentially expressed between MS and healthy subjects (CTRL), among which 629 were up-regulated and 838 were down-regulated (Fig. 1D, E).

3.3 WGCNA and identification of modules in MS

WGCNA was employed to identify the key modules in MS. Firstly, the optimal soft threshold was determined for making the created network more compatible with the scale-free topology. As shown in Fig. 2A, B, the scale-free topology fit index approached 0.85 when the power of the soft threshold was set to eight, as indicated by both scale independence and mean connectivity. Consequently, a soft thresholding value of eight was selected for subsequent analysis. Using this optimal threshold, we constructed a co-expression network and divided genes into different modules, generating a clustering dendrogram for the MS and control groups (Fig. 2C). A total of eight co-expression modules were identified and labeled with different colors according to the clustering dendrogram (Fig. 2D, E). Further analysis via heatmap revealed that the blue module, comprising 2481 genes, exhibited the most significant correlation with MS (correlation coefficient

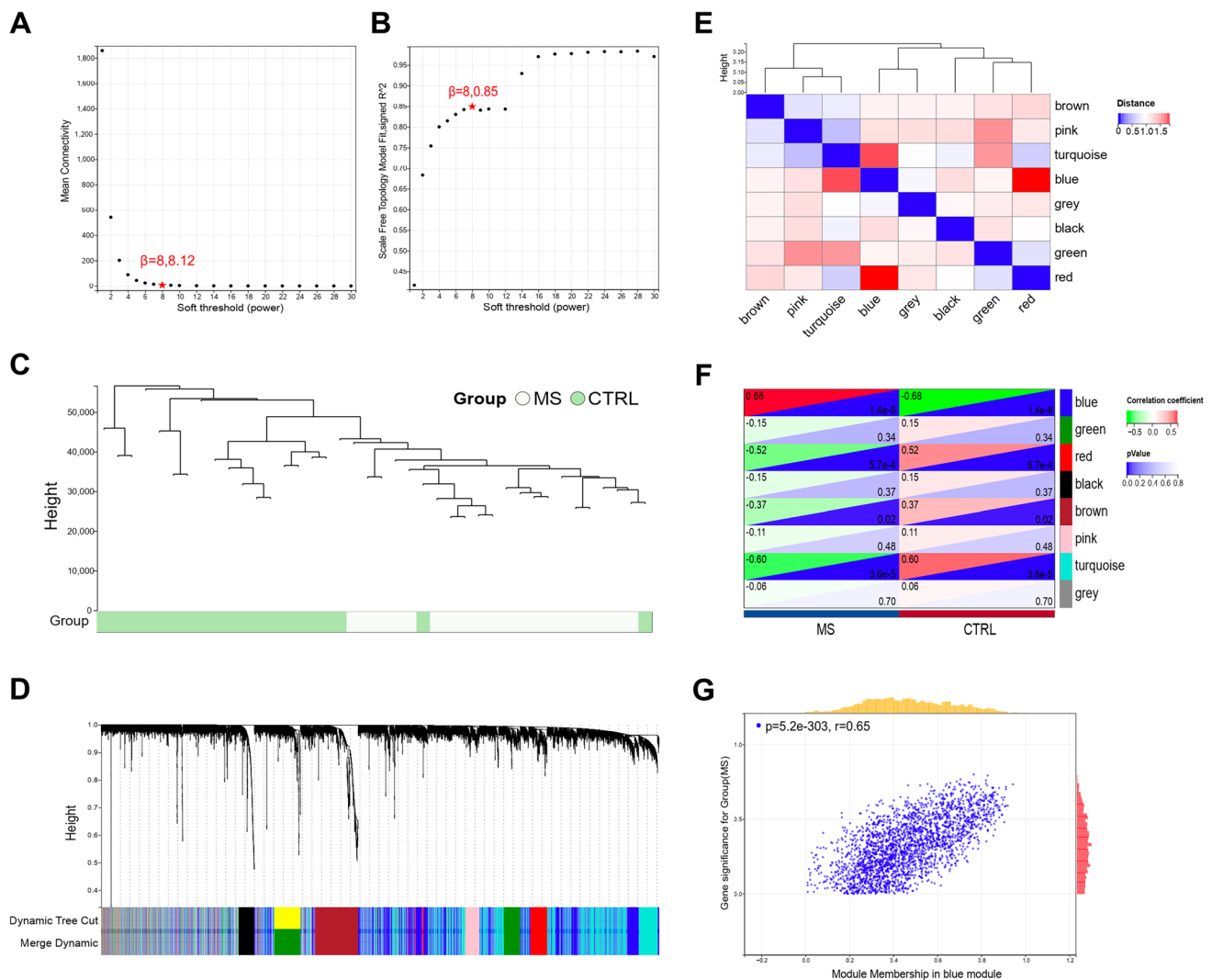


Fig. 2 Identification of DEGs and important module genes in MS patients. **A, B** Determination of soft thresholds based on criteria of scale-free networks. **C** Hierarchical clustering of module eigengenes. **D** Co-expression modules under the gene tree. **E** Heatmap showing the correlations between module eigengenes and clinical traits. **F** Association between the identified module and MS. **G** Association between module membership in the blue module and gene relevance for the MS group

$=0.68$, $P=1.4 \times 10^{-6}$), thus being identified as the key module for subsequent analysis (Fig. 2F). Additionally, a correlation plot depicted a strong positive association between the genes within the blue module and MS ($r=0.65$) (Fig. 2G).

3.4 Functional enrichment analysis of MS

Limma and WGCNA analyses were performed on the GSE98895 dataset, followed by GO and KEGG enrichment analyses, to investigate the potential pathogenesis of MS.

A Venn diagram revealed 276 intersecting genes between the Limma (1467 DEGs) and WGCNA (2481 genes) analyses (Fig. 3A). KEGG analysis indicated that these intersecting genes were primarily enriched in immune regulation genes, particularly pathways closely associated with T cell and NK cell functions as well as antigen presentation (Fig. 3B), suggesting a strong correlation between the development of MS and immune regulation. Similarly, the GO enrichment analysis of biological process (BP) terms also identified pathways related to immune regulation, such as lymphocyte activation, differentiation, and NK cell cytotoxicity, further indicating that MS is closely linked to immune modulation (Fig. 3C). In the cellular components (CC) terms of the GO enrichment analysis, the AP-1 adaptor complex was identified, a family associated with T cell activation and exhaustion (Fig. 3D). Notably, several antigen presentation-related pathways

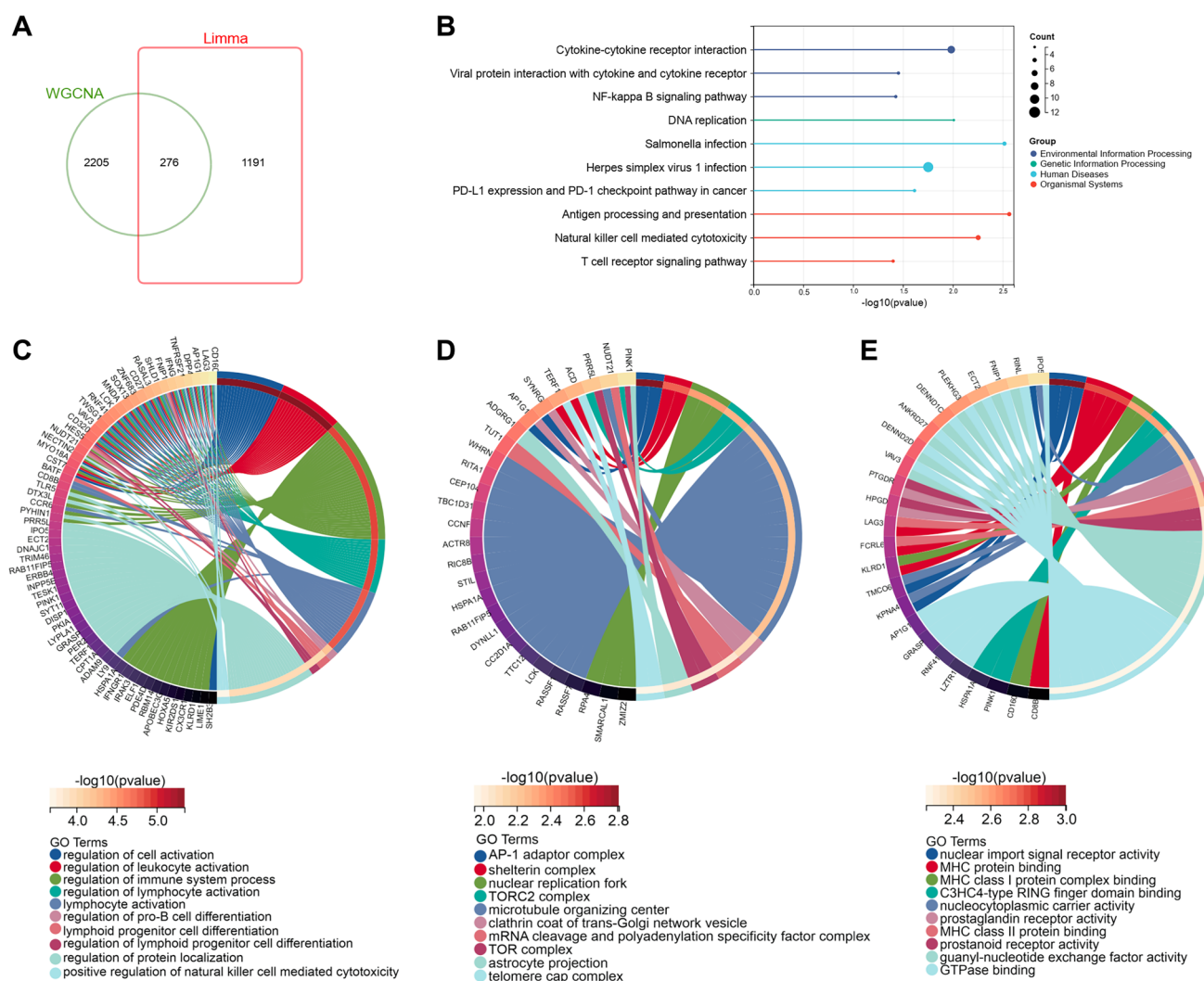


Fig. 3 Functional enrichment analysis for the intersections of WGCNA important module genes and DEGs in the MS database. **A** Venn diagram illustrating the intersecting genes between the important module genes identified by WGCNA and DEGs identified by Limma. **B** KEGG pathway enrichment of the intersecting genes. **C–E** GO enrichment of the intersecting genes, including biological processes (**C**), cellular components (**D**), and molecular functions (**E**)

were found among the enriched molecular functions (MF) terms, further confirming the role of immune-related pathways in the pathogenesis of MS (Fig. 3E).

3.5 Enrichment and PPI network of the DEGs in PCa-MS

We subsequently investigated the relevance of MS-related genes to the development of PCa. As illustrated in Fig. 4A, a total of 52 intersecting genes were distinguished between PCa DEGs and MS modular genes. The KEGG enrichment analysis of these intersecting genes revealed significant enrichment in pathways related to cell adhesion molecules, RNA polymerase, and the cytosolic DNA-sensing pathway (Fig. 4B). GO enrichment analysis indicated that these intersecting genes were associated with BP terms related to transcription and translational regulation, CC terms including the endoplasmic reticulum lumen and extracellular exosome, as well as molecular function MF terms such as transferase activity and kinase activity (Fig. 4C–E).

Based on functional enrichment analysis, we identified a significant association between the development of PCa-MS and immunomodulation. Subsequently, a PPI network was generated using the STRING database to investigate the interactions of DEGs in PCa-MS. The resulting PPI network, which comprises 51 nodes and 5 edges according to the

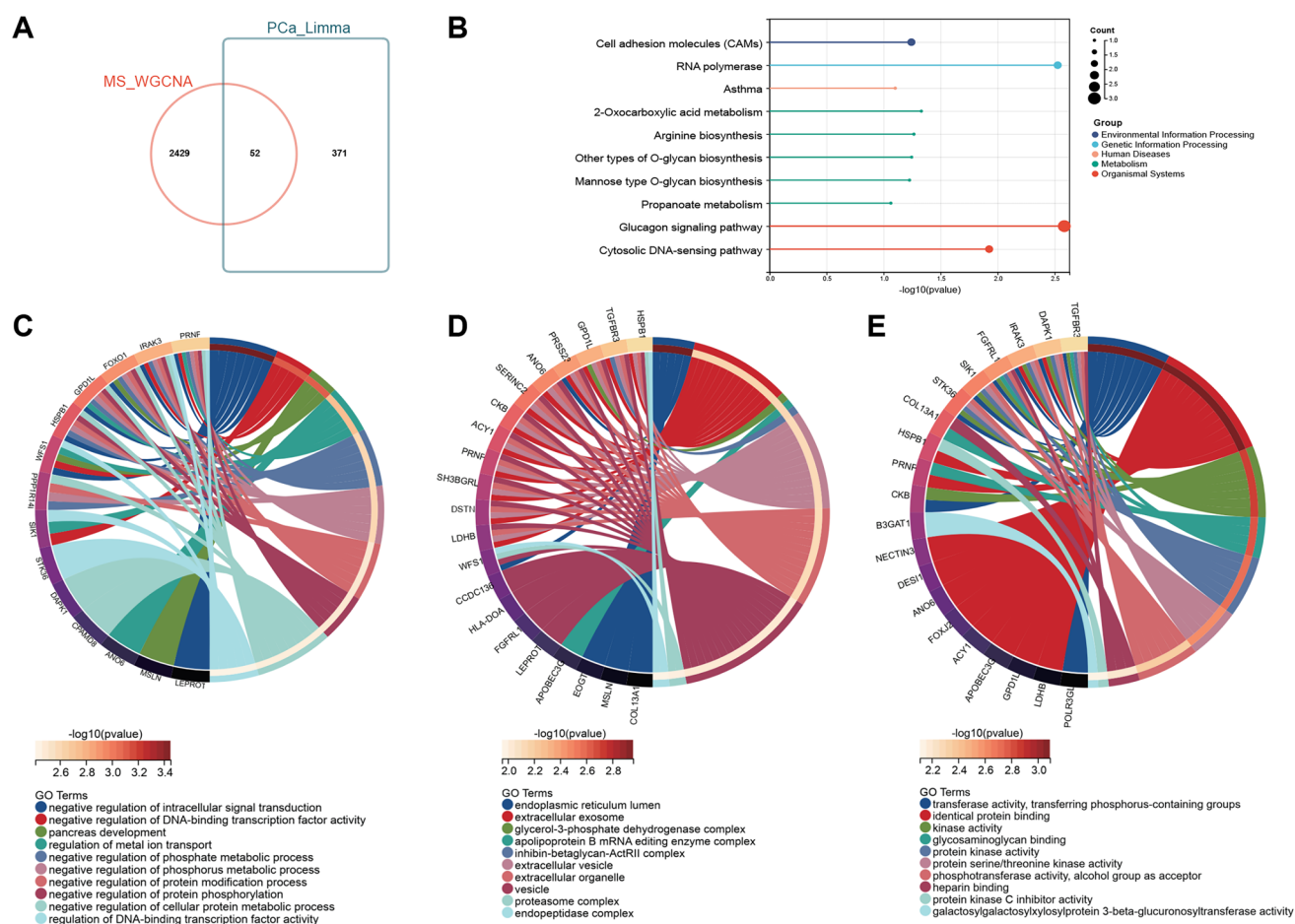


Fig. 4 Functional enrichment analysis for intersecting genes identified in PCa and MS patients. **A** Venn diagram illustrating the intersecting genes between DEGs identified in merged PCa databases by Limma and important module genes identified in the MS database by WGCNA. **B** KEGG pathway enrichment of the intersecting genes. **C–E** GO enrichment of the intersecting genes, including biological processes (**C**), cellular components (**D**), and molecular functions (**E**)

STRING database, was structured (Supplementary Fig. 2A). Essential modules were identified utilizing the MCODE plugin, and the genes were sorted based on their node degree (Supplementary Fig. 2B).

3.6 Identification of the signature genes for PCa-MS and analysis of their prognostic significance

Two machine learning algorithms were used to identify signature genes from the pool of 52 genes in PCa-MS. The LASSO regression screened 21 significant genes associated with PCa-MS (Fig. 5A, B). Meanwhile, the RF algorithm identified six signature genes with relative importance scores exceeding one (Fig. 5C, D). Subsequent crossover analysis highlighted the presence of three common signature genes between the LASSO analysis and the RF algorithm, namely ACY1, C12orf75, and GPD1L (Fig. 5E, Supplementary Fig. 3A, B). ROC (Fig. 5F) and precision recall curve (PRC) (Fig. 5G) analyses were conducted to assess the diagnostic specificity and sensitivity of each gene for PCa-MS. The AUC values were 0.85, 0.93 and 0.87 for ACY1, C12orf75 and GPD1L, respectively, indicating the high diagnostic value of these significant genes for PCa-MS.

Subsequently, through IHC staining, we observed differential expression of the aforementioned three proteins in clinical samples from our center. Compared to benign prostatic hyperplasia, ACY1 was found to be highly expressed in PCa, while the expression of C12orf75 and GPD1L was decreased in PCa (Fig. 5H). This was consistent with their expression trends at the transcriptional level (Supplementary Fig. 3A).

In addition, we analyzed the prognostic significance of these signature genes in PCa based on the TCGA database (cut-off establishment shown in Supplementary Fig. 4). We found that patients with high expression of ACY1 exhibited shorter PFS ($P = 0.0468$) and OS ($P = 0.0086$), suggesting its adverse prognostic significance. However, the

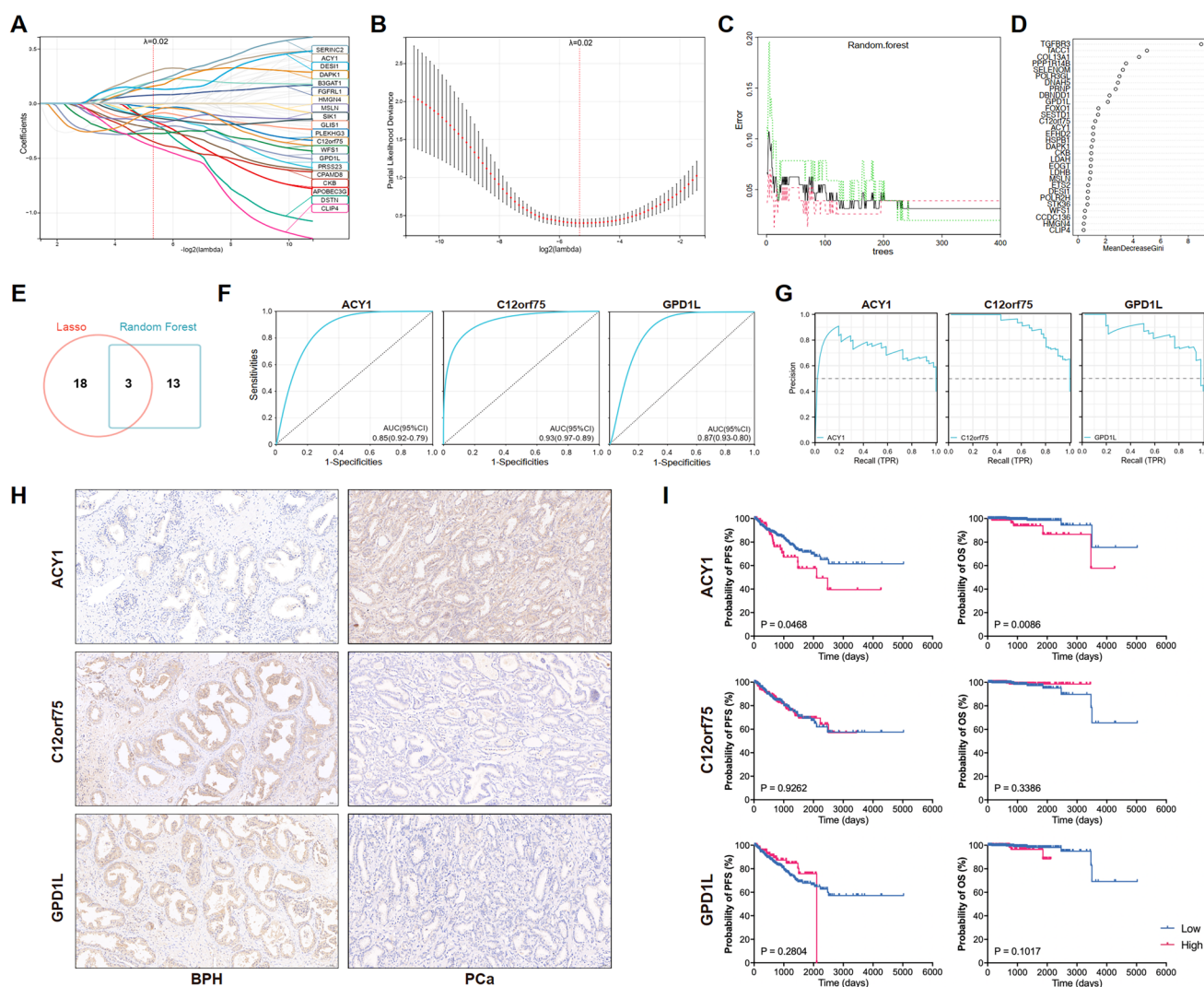


Fig. 5 Identification of the signature genes in PCa and MS. **A, B** LASSO coefficient profiles of the 21 significant genes. **C** Association between random forest error rate and a random number of decision trees. **D** Gene ranking based on importance scores derived from the random forest algorithm. **E** Venn diagram of common signature genes identified by both LASSO and random forest algorithm. **F** Receiver operating characteristic (ROC) curves of the three signature genes. **G** Precision recall curve (PRC) of the three signature genes. **H** Representation images of immunohistochemistry staining for the three signature genes. **I** Prognostic significance of the three signature genes

expression levels of C12orf75 and GPD1L showed no significant impact on the survival of PCa patients (Fig. 5I). The clinical characteristics of patients with different levels of ACY1 expression are presented in Supplementary Table 2.

3.7 Analysis of immune infiltration

To gain deeper insights into the relationship between PCa and immune regulation, we used CIBERSORT to analyze the immunological infiltration of 22 immune cell subsets in the merged PCa databases. By examining the percentages of these 22 immune cells in the PCa and BPH groups (Fig. 6A), we found significant differences ($P < 0.05$) in infiltration between the control and PCa groups, involving B-cell memory, T-cell gamma delta, macrophage M0, as well as resting and activated dendritic cells (DCs) (Fig. 6B). The correlation heatmap indicated that B cell naïve, activated DCs, and activated mast cells were negatively correlated with B cell memory, M0 macrophages, and resting mast cells, respectively. Additionally, a moderate correlation was observed between CD8 + T cells and neutrophils (Fig. 6C).

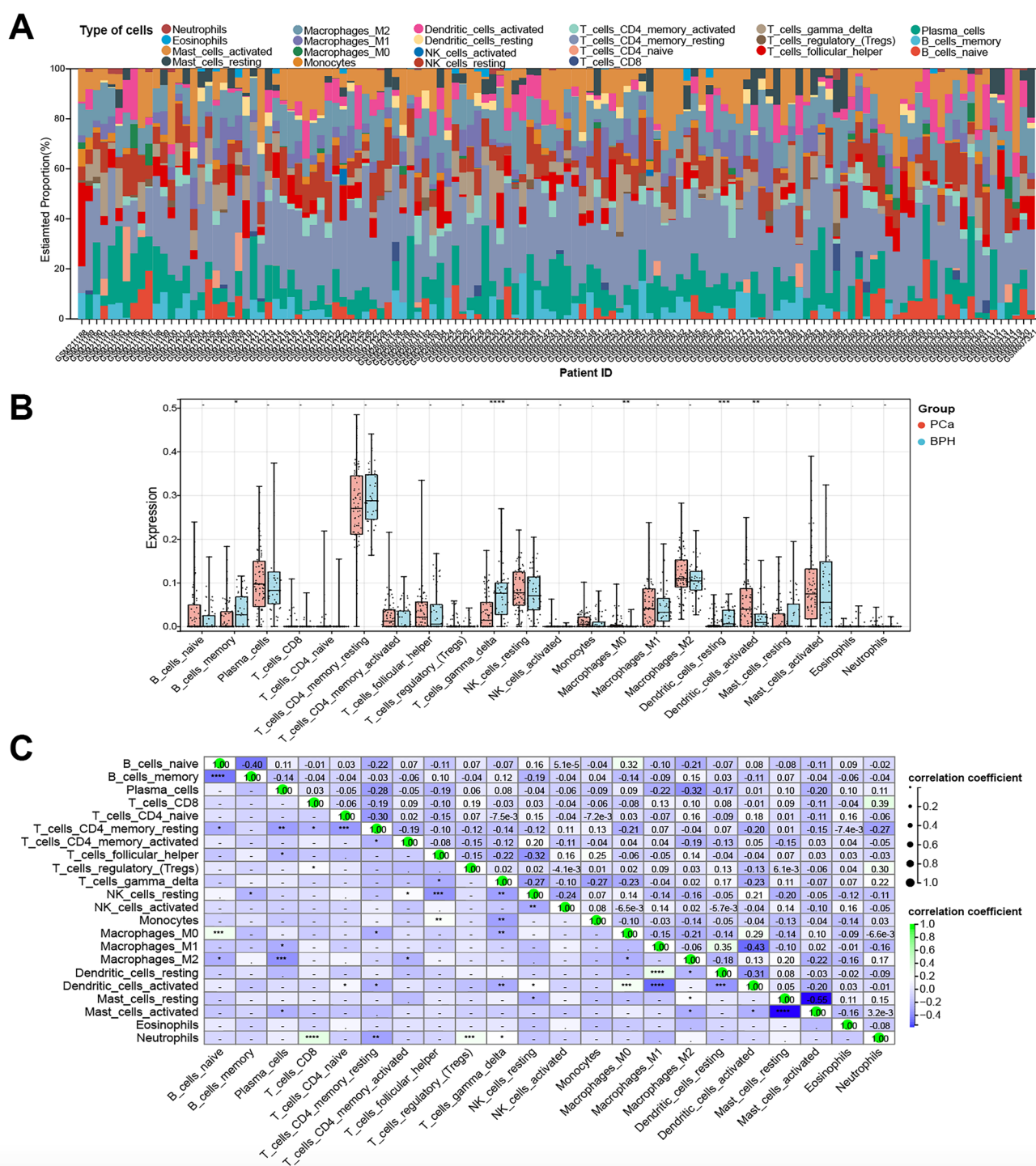


Fig. 6 Immune infiltration analysis for the merged PCa database. **A** Relative abundance of 22 immune cell types in different samples. **B** Histogram showing the ratio of 22 immune cell types in PCa and BPH groups. **C** Heatmap of correlated among the 22 types of immune cells.

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

4 Discussion

PCa is one of the most common malignancies affecting men and poses a significant threat to male life expectancy. MS is also increasingly recognized as a pathological condition that affects quality of life and can lead to fatal complications

in middle-aged and elderly individuals. Interestingly, PCa, being an endocrine -related tumor, is closely associated with metabolic processes. Some studies suggest that MS may increase the risk of PCa [20, 21], while others indicate no significant impact on PCa incidence [22, 23], though associations with advanced and recurrent PCa have been noted [24], contributing to higher mortality rates even when considering the competing risk of mortality [22]. Additionally, treatment for PCa can induce MS, with both conditions sharing a chronic inflammation microenvironment. Therefore, a deeper understanding of the mechanisms underlying the development of both conditions, especially the genes or pathways involved, may provide new perspectives and potential targets for treating PCa, particularly PCa-MS.

During the functional enrichment analysis of the MS database, we observed that the DEGs between MS patients and non-MS individuals are significantly enriched in immune regulatory pathways, particularly those involving lymphocytes and NK cell regulation. Previous studies have demonstrated that pro-inflammatory factors in adipose tissue recruit immune cells such as CD4 + T cells, NK cells, and innate lymphoid cells, exacerbating inflammation and contributing to MS [25, 26]. Our findings further support the link between immune dysregulation and MS pathogenesis. Subsequently, we analyzed the common DEGs in both MS and PCa to explore shared pathogenic background. The results indicated that transcriptional and translational regulation is closely related to both diseases. This finding aligns with the role of the key gene ACY1, which is involved in protein catabolism. Recent reports have highlighted the critical role of translational regulation in both MS [27, 28] and prostate cancer [29]. Recent advances in proteomics offer potential for further elucidating shared pathogenic backgrounds and identifying therapeutic targets.

Using comprehensive bioinformatics and machine learning methods, we identified ACY1, C12orf75, and GPD1L as signature genes for PCa-MS. Clinical samples from our center confirmed ACY1 is upregulated in PCa, while C12orf75 and GPD1L are downregulated compared to benign prostatic hyperplasia. These findings suggest an association of these genes with the onset of both PCa and MS. Notably, our subsequent survival analysis revealed that high ACY1 expression is associated with unfavorable prognosis.

Aminoacylase 1 (ACY1) is a homodimeric zinc-binding metalloenzyme crucial for deacylation of α -acylated amino acid residues from the N-terminal peptide during protein catabolism [30]. ACY1 has been reported to be involved in tumorigenesis, with its expression and prognostic significance vary across cancers, being increased in colorectal cancer [31], rectal carcinoma [32], and intrahepatic cholangiocarcinoma [33], but decreased in renal cell carcinoma [34] and hepatocellular carcinoma [35]. In colorectal cancer [36] and non-small cell lung cancer [37], high ACY1 expression predicts poor prognosis, whereas low expression correlates with poor prognosis in neuroblastoma [38]. In this study, we observed upregulation of ACY1 in PCa through database analysis and clinical specimen examination, further indicating its association with poor prognosis. These findings suggest that ACY1 holds potential as both a therapeutic target and a prognostic marker in PCa. We propose that ACY1 may contribute to tumorigenesis by modulating amino acid metabolism, supporting cellular detoxification, and maintaining cellular homeostasis.

C12orf75 and GPD1L are also considered closely associated with tumor initiation and progression. Glycerol-3-phosphate dehydrogenase 1-like (GPD1L) is a protease that catalyzes the conversion of glycerol triphosphate to glycerol phosphorene. As a novel regulator of HIF1 α stability, GPD1L plays a role in attenuating the hypoxic response and regulating tumor growth in various types of tumors [39, 40]. Chromosome 12 open reading frame 75 (C12orf75), also named OCC-1, has been reported to regulate tumor growth in colorectal cancer [41] and breast cancer [42]. Reports on these genes in PCa and MS are limited, and their roles in tumors remain unclear. Our analysis did not reveal the prognostic significance of these genes in PCa, highlighting tumor heterogeneity and microenvironment complexity. Therefore, further exploration of these genes in PCa and MS may provide valuable insights for a more detailed understanding of their roles.

Previous research has shown a strong connection between immune modulation and the development and progression of PCa [43]. We found elevated macrophage M0 and dendritic cell infiltration in PCa tissues, consistent with previous studies [43], suggesting these immune cells as potential diagnosis and therapeutic targets. Moreover, lower resting levels of B-cell memory, T-cell gamma delta, and dendritic cells were observed in PCa compared to the control group.

5 Conclusions

To conclude, by employing advanced bioinformatics and machine learning, we elucidate the genetic characteristics of PCa and MS, identifying shared signature genes for both conditions. With the increasing incidence rates of PCa and MS, especially given their interconnected nature and frequent co-occurrence, our study offers novel insights into

potential biomarkers and the immune landscape of PCa and MS, providing new perspectives for exploring potential targets.

Acknowledgements We thank the patients and their families for their participation. We appreciate Dr. Kaiyan Chen for the excellent technical support in this study.

Author contributions Y.X., and M.G. designed the research. Y.X., M.G., and S.L. analyzed the data of the datasets. M.G., S.L., X.J., and Z.X. collected the clinical samples and performed the assays. M.G., Z.X., Z.G., J.M., W.Z., and H.B. collected and analyzed the clinical data. M.G., Y.X., X.J., J.Z., and J.W. interpreted and visualize the data. M.G. drafted the manuscript, Y.X. revised the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by the Program of Taizhou School of Clinical Medicine of Nanjing Medical University (NO. TZKY20230106, NO. TZKY20220310, and No. TZKY20240201), Research Program of Taizhou People's Hospital (No. ZL202203), and Natural Science Foundation of Jiangsu Province (No. BK20241119).

Data availability Publicly available datasets were analyzed in this study. These data can be found at <https://www.ncbi.nlm.nih.gov/geo/> (GSE8511, GSE32571, GSE104749 and GSE98895) and <https://portal.gdc.cancer.gov/> (TCGA). The code used for analysis and generating the figures in the study can be obtained by contacting the corresponding author.

Declarations

Ethics approval and consent to participate This study was conducted in accordance with the Declaration of Helsinki, and was approved by the institutional review boards of the Affiliated Taizhou People's Hospital of Nanjing Medical University Ethics Committee (Ethics approval number: KY2023-011–01). Informed consent from patients was obtained for tissue use.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

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/>.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–49.
2. Siegel RL, Miller KD, Wagie NS, et al. Cancer statistics, 2023. *CA Cancer J Clin.* 2023;73(1):17–48.
3. Sung H, Siegel RL, Torre LA, et al. Global patterns in excess body weight and the associated cancer burden. *CA Cancer J Clin.* 2019;69(2):88–112.
4. Lemieux I, Despres JP. Metabolic syndrome: past, present and future. *Nutrients.* 2020;12(11):3501.
5. Gandaglia G, Leni R, Bray F, et al. Epidemiology and prevention of prostate cancer. *Eur Urol Oncol.* 2021;4(6):877–92.
6. Hernandez-Perez JG, Torres-Sanchez L, Hernandez-Alcaraz C, et al. Metabolic syndrome and prostate cancer risk: a population case–control study. *Arch Med Res.* 2022;53(6):594–602.
7. Lifshitz K, Ber Y, Margel D. Role of metabolic syndrome in prostate cancer development. *Eur Urol Focus.* 2021;7(3):508–12.
8. He Y, Xu W, Xiao YT, et al. Targeting signaling pathways in prostate cancer: mechanisms and clinical trials. *Signal Transduct Target Ther.* 2022;7(1):198.
9. Andersen CJ, Murphy KE, Fernandez ML. Impact of obesity and metabolic syndrome on immunity. *Adv Nutr.* 2016;7(1):66–75.
10. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* 2007;8(1):118–27.
11. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7): e47.
12. Zhou Y, Shi W, Zhao D, et al. Identification of immune-associated genes in diagnosing aortic valve calcification with metabolic syndrome by integrated bioinformatics analysis and machine learning. *Front Immunol.* 2022;13:937886.
13. Dai B, Bailey-Kellogg C. Protein interaction interface region prediction by geometric deep learning. *Bioinformatics.* 2021;37(17):2580–8.
14. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (N Y).* 2021;2(3):100141.
15. Chen D, Liu J, Zang L, et al. Integrated machine learning and bioinformatic analyses constructed a novel stemness-related classifier to predict prognosis and immunotherapy responses for hepatocellular carcinoma patients. *Int J Biol Sci.* 2022;18(1):360–73.

16. Zhao Z, Yang YB, Li XY, et al. Comprehensive analysis of N6-methyladenosine-related lncRNA signature for predicting prognosis and immune cell infiltration in patients with colorectal cancer. *Dis Markers*. 2021;2021:8686307.
17. Jorgenson C, Higgins O, Petrelli M, et al. A machine learning-based approach to clinopyroxene thermobarometry: model optimization and distribution for use in earth sciences. *J Geophys Res Solid Earth*. 2022;127(4):e2021JB022904.
18. Taminau J, Meganck S, Lazar C, et al. Unlocking the potential of publicly available microarray data using inSilicoDb and inSilicoMerging R/Bioconductor packages. *BMC Bioinform*. 2012;13:335.
19. Yang Y, Liu P, Teng R, et al. Integrative bioinformatics analysis of potential therapeutic targets and immune infiltration characteristics in dilated cardiomyopathy. *Ann Transl Med*. 2022;10(6):348.
20. Yoo S, Oh S, Park J, et al. Effects of metabolic syndrome on the prevalence of prostate cancer: historical cohort study using the national health insurance service database. *J Cancer Res Clin Oncol*. 2019;145(3):775–80.
21. Long XJ, Lin S, Sun YN, et al. Diabetes mellitus and prostate cancer risk in Asian countries: a meta-analysis. *Asian Pac J Cancer Prev*. 2012;13(8):4097–100.
22. Grundmark B, Garmo H, Loda M, et al. The metabolic syndrome and the risk of prostate cancer under competing risks of death from other causes. *Cancer Epidemiol Biomarkers Prev*. 2010;19(8):2088–96.
23. Osaki Y, Taniguchi S, Tahara A, et al. Metabolic syndrome and incidence of liver and breast cancers in Japan. *Cancer Epidemiol*. 2012;36(2):141–7.
24. Burton AJ, Gilbert R, Tilling K, et al. Circulating adiponectin and leptin and risk of overall and aggressive prostate cancer: a systematic review and meta-analysis. *Sci Rep*. 2021;11(1):320.
25. Saitoh S, Van Wijk K, Nakajima O. Crosstalk between metabolic disorders and immune cells. *Int J Mol Sci*. 2021;22(18):10017.
26. Lackey DE, Olefsky JM. Regulation of metabolism by the innate immune system. *Nat Rev Endocrinol*. 2016;12(1):15–28.
27. Russo S, Kwiatkowski M, Govorukhina N, et al. Meta-inflammation and metabolic reprogramming of macrophages in diabetes and obesity: the importance of metabolites. *Front Immunol*. 2021;12:746151.
28. Kitamura H. Ubiquitin-specific proteases (USPs) and metabolic disorders. *Int J Mol Sci*. 2023;24(4):3219.
29. Intasqui P, Bertolla RP, Sadi MV. Prostate cancer proteomics: clinically useful protein biomarkers and future perspectives. *Expert Rev Proteomics*. 2018;15(1):65–79.
30. Goverti D, Yuksel RN, Kaya H, et al. Serum concentrations of aminoacylase 1 in schizophrenia as a potential biomarker: a case-sibling-control study. *Nord J Psychiatry*. 2022;76(5):380–5.
31. Yu B, Liu X, Cao X, et al. Study of the expression and function of ACY1 in patients with colorectal cancer. *Oncol Lett*. 2017;13(4):2459–64.
32. Xu Z, Hu Y, Yu Z. Effect of the ACY-1 gene on HER2 and TRAIL expression in rectal carcinoma. *Exp Ther Med*. 2021;22(2):817.
33. Cavalloni G, Peraldo-Neia C, Massa A, et al. Proteomic analysis identifies deregulated metabolic and oxidative-associated proteins in Italian intrahepatic cholangiocarcinoma patients. *BMC Cancer*. 2021;21(1):865.
34. Zhong Y, Onuki J, Yamasaki T, et al. Genome-wide analysis identifies a tumor suppressor role for aminoacylase 1 in iron-induced rat renal cell carcinoma. *Carcinogenesis*. 2009;30(1):158–64.
35. Wei X, Li J, Xie H, et al. Proteomics-based identification of the tumor suppressor role of aminoacylase 1 in hepatocellular carcinoma. *Cancer Lett*. 2014;351(1):117–25.
36. Shi H, Hayes MT, Kirana C, et al. Overexpression of aminoacylase 1 is associated with colorectal cancer progression. *Hum Pathol*. 2013;44(6):1089–97.
37. Chen H, Wang W, Xiao C, et al. ACY1 regulating PTEN/PI3K/AKT signaling in the promotion of non-small cell lung cancer progression. *Ann Transl Med*. 2021;9(17):1378.
38. Long PM, Stradecki HM, Minturn JE, et al. Differential aminoacylase expression in neuroblastoma. *Int J Cancer*. 2011;129(6):1322–30.
39. Liu H, Wang S, Cheng A, et al. GPD1L is negatively associated with HIF1alpha expression and predicts lymph node metastasis in oral and HPV-oro-pharyngeal cancer. *Oral Dis*. 2021;27(7):1654–66.
40. Kelly TJ, Souza AL, Clish CB, et al. A hypoxia-induced positive feedback loop promotes hypoxia-inducible factor 1alpha stability through miR-210 suppression of glycerol-3-phosphate dehydrogenase 1-like. *Mol Cell Biol*. 2011;31(13):2696–706.
41. Najafi H, Soltani BM, Dokanehiifard S, et al. Alternative splicing of the OCC-1 gene generates three splice variants and a novel exonic microRNA, which regulate the Wnt signaling pathway. *RNA*. 2017;23(1):70–85.
42. Ghalei A, Kay M, Zarrinfam S, et al. Overexpressed in colorectal carcinoma gene (OCC-1) upregulation and APPL2 gene downregulation in breast cancer specimens. *Mol Biol Rep*. 2018;45(6):1889–95.
43. Wu Z, Chen H, Luo W, et al. The Landscape of Immune Cells Infiltrating in Prostate Cancer. *Front Oncol*. 2020;10:517637.

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