Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Machine Learning of Tumor Microenvironment Related Genes for Prostate Cancer Biochemical Recurrence Prediction and Immune Analysis

Taize Sun
Hainan General Hospital, Affiliated Hainan Hospital of Hainan Medical University
JunCheng Chen
Hainan General Hospital, Affiliated Hainan Hospital of Hainan Medical University
Fei Wang
hnsywangfei@163.com

Hainan General Hospital, Affiliated Hainan Hospital of Hainan Medical University

Research Article

Keywords: Prostate cancer, biochemical recurrence, machine learning, prognostic model, tumor microenvironment, drug sensitivity

Posted Date: June 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9645313/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Machine Learning of Tumor Microenvironment Related Genes for Prostate Cancer Biochemical Recurrence Prediction and Immune Analysis

Taize Sun[1,2,†], JunCheng Chen[1,2], Fei Wang[1,2#]

[1]Department of Urology, Hainan General Hospital, Affiliated Hainan Hospital of Hainan Medical University, Hainan Provincial Clinical Medical Center, Haikou 570311, China

[2]School of Clinical Medicine, Hainan Medical University, Haikou, Hainan 571199, China

[†]These authors contributed equally to this work.

[#]Correspondence address. Tel: 13398996987; E-mail: hnsywangfei@163.com

Abstract

Background: Biochemical recurrence (BCR) remains a major clinical challenge after radical prostatectomy in prostate cancer (PCa). However, the prognostic value of tumor microenvironment (TME)-related genes in BCR has not been fully elucidated. This study aimed to construct and validate a TME-related prognostic model for BCR in PCa and to explore its associations with immune infiltration, immunotherapy response, and individualized prognosis assessment.

Methods: Transcriptomic and clinical data from the TCGA and GEO cohorts were integrated. The ESTIMATE algorithm was applied to identify TME-related differentially expressed genes that were commonly associated with both ImmuneScore and StromalScore. Univariate Cox regression analysis was performed to screen candidate genes significantly correlated with BCR. An optimal prognostic model was then constructed using 117 machine learning algorithm combinations. Model performance was evaluated using Kaplan–Meier analysis, time-dependent ROC curves, the concordance index (C-index), and external validation cohorts. SHAP analysis was used to interpret the contribution of core genes. Additional analyses included immune profiling, somatic mutation

landscape, drug sensitivity, weighted gene co-expression network analysis (WGCNA), single-cell sequencing, and in silico knockout experiments.

Results: A total of 32 TME-related genes were significantly associated with BCR, and a robust optimal prediction model was established. SHAP analysis identified five core contributing genes. The high-risk group exhibited higher immune scores, upregulation of immunosuppressive components and most immune modulatory molecules, as well as an elevated TIDE score, indicating a stronger propensity for immune escape. WGCNA revealed key modules primarily enriched in extracellular matrix organization and collagen metabolism. Drug sensitivity analysis suggested that the high-risk group might be more sensitive to oxaliplatin and cyclophosphamide.

Conclusion: A robust BCR prediction model based on TME-related genes was successfully constructed and validated for PCa. This model can be used for postoperative BCR risk stratification and individualized prognosis assessment, and provides a basis for elucidating immune regulatory mechanisms and optimizing treatment strategies.

Keywords:Prostate cancer; biochemical recurrence; machine learning; prognostic model; tumor microenvironment; drug sensitivity

Introduction

PCa ranks among the most prevalent malignancies in middle-aged and elderly men[1]. For the majority of patients with localized PCa, radical prostatectomy (RP) remains the primary curative intervention[2]. However, due to the heterogeneous biological behavior of PCa—characterized by the coexistence of indolent and highly aggressive phenotypes—BCR occurs in approximately 25% of patients within four years following RP[3-4].

BCR represents a significant clinical challenge in the postoperative management of PCa. Once

BCR develops, a substantial proportion of patients will inevitably progress to castration-resistant prostate cancer, a transition that markedly worsens prognosis[5-6]. Consequently, the establishment of reliable risk stratification tools for BCR and an in-depth exploration of the underlying molecular mechanisms are of paramount importance for optimizing individualized therapeutic strategies for patients with PCa

.Accumulating evidence indicates that tumor initiation, progression, and therapeutic response are governed not solely by the genetic and epigenetic aberrations intrinsic to cancer cells but also by dynamic regulation within the TME. The interplay between malignant cells and surrounding supportive components fosters the acquisition of aggressive malignant traits, including invasive metastasis and immune evasion[7]. The TME is predominantly composed of resident stromal cells and recruited immune cells. Within this milieu, cancer-associated fibroblasts (CAFs), endothelial cells, and pericytes play pivotal roles in orchestrating tumor angiogenesis[8-10], while CAFs and adipocytes actively participate in extracellular matrix remodeling and the facilitation of epithelial-mesenchymal transition[11-13]. Moreover, the immunomodulatory functions exerted by immune cells within the TME on tumor growth and progression have garnered considerable attention. In the context of PCa, the TME frequently manifests pronounced immunosuppressive features, typified by increased heterogeneity of immune cell infiltration, insufficient antigen presentation, and restricted recruitment and activation of cytotoxic T lymphocytes[14-16].

Prior investigations have corroborated a close association between the infiltration status of various immune cell subsets within the TME and BCR. Notably, tumor-associated macrophage infiltration has been identified as a predictive factor for BCR[17], and B cell infiltration in PCa has similarly been demonstrated to correlate positively with BCR occurrence[18]. Aberrant regulation of stromal and immune constituents within the TME may represent a critical precipitating factor driving the formation of micrometastatic foci that ultimately culminate in BCR. Accordingly, a focused examination of TME dynamics may facilitate the identification of potential molecular biomarkers for predicting BCR risk.

With the advent of high-throughput sequencing technologies and integrative multi-omics analyses, numerous prognostic models predicated on gene expression profiles have been developed for predicting BCR in PCa patients[19-20]. Nonetheless, existing studies are constrained by a limited exploration of TME-relevant genes[21]. Although certain investigations have incorporated TME-associated genes, they frequently rely on isolated algorithms or a restricted set of candidate

molecules, thereby lacking a comprehensive, integrative analysis encompassing regulatory pathways, prognostic significance, and immune microenvironmental landscape. Furthermore, given that the selection of machine learning algorithms can profoundly influence model stability and generalizability, findings derived from a singular modeling approach are susceptible to inherent bias. Thus, a multi-algorithm integration strategy is imperative for identifying the optimal model to enhance the clinical applicability of prognostic prediction.

In the present study, the ESTIMATE algorithm was employed to compute the immune and stromal scores of tumor samples, thereby facilitating the identification of differentially expressed genes within the TME. A large-scale machine learning-based computational framework was subsequently implemented to select the optimal prognostic model. Through SHAP-based interpretability analysis, characterization of clinical features across distinct risk subgroups, comprehensive dissection of the immune microenvironment composition, mutational landscape profiling, drug sensitivity prediction, single-cell sequencing analysis, and in silico gene perturbation experiments, the underlying mechanisms are elucidated from the perspective of immune-stromal crosstalk. This work provides a novel theoretical foundation for the precise prognostic assessment and the optimization of therapeutic strategies for patients with PCa.

1. Methods

1.1 Data acquisition and preprocessing

The training cohort comprised PCa samples with complete BCR information obtained from The Cancer Genome Atlas (TCGA) database (n = 408). For external validation, the microarray datasets GSE70769 (n = 92), GSE46602 (n = 36), and GSE116918 (n = 248) were retrieved from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). Microarray data were normalized using the normalizeBetweenArrays function within the R package limma, whereas RNA sequencing data underwent transcripts per million (TPM) transformation followed by $\log_2$ transformation. Batch effects were corrected to ensure cross-platform data comparability.

1.2 Calculation of TME-associated scores and screening of differentially expressed genes

The ESTIMATE algorithm was employed to assess the abundance of immune and stromal components in TCGA-PCa samples, yielding ImmuneScore, StromalScore, and ESTIMATEScore for each specimen. Samples were subsequently stratified into high- and low-score groups based on the median values of ImmuneScore and StromalScore, respectively. Differential expression analysis

between the high- and low-score groups was performed using the limma package, with the screening criteria set at an adjusted P-value < 0.05 and $|\log_2$ fold change (FC)| > 1. The intersection of differentially expressed genes (DEGs) derived from the ImmuneScore and StromalScore comparisons was subsequently identified, generating a set of common DEGs designated as TME-associated candidate genes. The overlap was visualized using a Venn diagram. Protein–protein interaction (PPI) networks were constructed utilizing the STRING database and visualized with Cytoscape.

1.3 Functional enrichment analysis

To elucidate the potential biological functions of the TME-associated DEGs, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were conducted using the R package clusterProfiler. The significance threshold was set at $P < 0.05$. Concurrently, single-sample gene set enrichment analysis (ssGSEA) was performed on the entire expression matrix to identify functional pathways associated with TME status. For ssGSEA of single-cell data, the mean expression profile of each cell type served as the input. Enrichment results were visualized employing ggplot2 and related visualization utilities.

1.4 Identification of BCR-associated candidate genes

Based on the TME-associated DEGs obtained above, univariate Cox proportional hazards regression analysis was performed in conjunction with the BCR follow-up data of TCGA-PCa patients to identify candidate genes significantly associated with BCR. The screening criterion was set at $P < 0.05$. Genes meeting this threshold were retained for the subsequent construction of machine learning-based prognostic models.

1.5 Construction of the ensemble machine learning prognostic model

To establish a robust predictive model for BCR, an integrative machine learning framework was constructed using the R package MIME[22]. Specifically, a total of 117 algorithm combinations were generated based on ten commonly employed survival analysis algorithms, including Lasso, Ridge, Elastic Net (Enet), CoxBoost, Stepwise Cox regression (StepCox), partial least squares Cox regression (plsRcox), supervised principal components analysis (SuperPC), generalized boosted regression modeling (GBM), random survival forest (RSF), and survival support vector machine (survival-SVM). Each model was fitted using the TCGA training cohort and subsequently evaluated in the external validation cohorts. The optimal model was identified by comparing the mean

concordance index (C-index) across both training and validation sets. Ultimately, the combination of CoxBoost and forward StepCox was determined to be the optimal model. Risk scores were calculated for each patient based on this model, and patients were stratified into high- and low-risk groups using the optimal cutoff value or median risk score.

1.6 Validation and performance assessment of the prognostic model

Kaplan–Meier survival analysis was employed to compare BCR-free survival differences between the high- and low-risk groups, with statistical significance assessed using the log-rank test. Time-dependent receiver operating characteristic (ROC) curves at 1-, 3-, and 5-year intervals were generated using the timeROC package, and the area under the curve (AUC) was calculated to evaluate the predictive capacity of the model. The C-index was further computed to assess overall model discrimination. The performance of the present model was compared with that of previously published prognostic models to validate its robustness and superiority.

1.7 SHAP-based interpretability analysis

To enhance model interpretability, SHapley Additive exPlanations (SHAP) analysis was performed on the feature genes incorporated within the optimal model. SHAP values were calculated for each gene to quantify both the direction and magnitude of their contributions to individual risk predictions. Core contributing genes were identified based on mean absolute SHAP values, and their influence on risk score patterns was visualized using SHAP summary plots and dependence plots.

1.8 Clinical correlation analysis and nomogram construction

To assess the clinical relevance of the risk score, its distribution was compared across different clinicopathological subgroups, including BCR status, Gleason score, PSA,T stage, N stage, and M stage. Univariate and multivariate Cox regression analyses were subsequently performed to identify independent prognostic factors. On this basis, a nomogram integrating the risk score and relevant clinical parameters was constructed to predict the probability of BCR at 1, 3, and 5 years. Calibration curves were generated to evaluate the predictive accuracy of the nomogram.

1.9 GSEA and immune infiltration analysis

To investigate the biological discrepancies between the high- and low-risk groups, GSEA was conducted to examine functional pathway alterations. Subsequently, the ESTIMATE algorithm was employed via the corresponding R package to compare ImmuneScore, StromalScore,

ESTIMATEScore, and tumor purity across the different risk strata[23]. The relative enrichment levels of various immune cell subpopulations in each risk group were assessed using ssGSEA. The CIBERSORT algorithm was utilized to estimate the relative proportions of 22 tumor-infiltrating immune cell types (TICs). Additionally, the xCell method was applied to further validate the compositional differences in immune and stromal cells between the risk groups. Finally, correlation analyses were performed between the expression of core model genes and immune cell infiltration levels to explore potential associations with the local immune milieu.

1.10 Analysis of immune checkpoint molecule expression and TIP analysis

To characterize immunotherapeutic features across different risk subgroups, the expression levels of multiple immune-regulatory molecules were compared between the high and low-risk groups. Visualizations were generated using the R packages ggplot2, ggpubr, and reshape2. Furthermore, the Tracking Tumor Immunophenotype (TIP) platform (http://biocc.hrbmu.edu.cn/TIP/) was utilized to evaluate the activity of the seven-step cancer-immunity cycle across risk subgroups. Differences in the anti-tumor immune process were assessed by comparing the activity scores of each step.

1.11 Mutational landscape analysis

Somatic mutation data for PCa patients were retrieved from the TCGA database. The mutational landscapes of the distinct risk subgroups were visualized and compared using the R package maftools. Comparative analysis of mutational profiles was conducted to explore potential relationships between the risk model and underlying genomic alterations.

1.12 Identification of key modules associated with risk and immunity via WGCNA

To further identify key gene modules associated with risk stratification and immune characteristics, DEGs between the high- and low-risk groups were first identified. Weighted gene co-expression network analysis (WGCNA) was subsequently performed based on these DEGs. An appropriate soft-thresholding power was selected to construct an approximate scale-free network. Co-expression modules were identified, and the correlations between module eigengenes, risk grouping, and immune scores were calculated. The module exhibiting significant correlations with both risk and immunity was selected for further investigation. Hub genes within this key module were identified based on module membership (MM) and gene significance (GS).

1.13 Functional enrichment and correlation analysis of hub genes

GO functional enrichment analysis and PPI network construction were performed for the hub genes identified through WGCNA to elucidate their potential biological functions. Concurrently, Spearman correlation analysis was employed to evaluate the relationships between hub gene expression, ESTIMATE scores, and immune cell infiltration levels, thereby exploring their potential roles in immune regulation and stromal remodeling.

1.14 Immunotherapy and drug sensitivity analysis

The R package oncoPredict was utilized to predict the potential sensitivity of patients in different risk groups to various anti-tumor drugs. The half-maximal inhibitory concentration (IC_{50}) for candidate drugs was estimated for each sample based on gene expression profiles. The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was employed to assess the potential for tumor immune evasion.

1.15 Single-cell RNA sequencing analysis

Single-cell RNA sequencing (scRNA-seq) data of primary PCa samples (GSE157703) were obtained from the GEO database and processed using the R package Seurat. Batch effects between samples were mitigated using the harmony package, and data normalization was performed via the ScaleData function. Principal component analysis (PCA) was subsequently conducted on the normalized data, followed by dimensionality reduction using the RunUMAP function. Differentially expressed genes within each cluster were identified using the FindAllMarkers function. Cell type annotation was accomplished using the R package scMayoMap. Finally, in silico knockout analyses of DPT and MNDA were performed using the R package scTenifoldKnk [24].

1.16 Statistical analysis

All statistical analyses were conducted using R software (version 4.4.3). Comparisons between groups were performed using the appropriate methods, including the *t*-test, Wilcoxon rank-sum test, or Kruskal–Wallis test, as dictated by the data distribution. Survival analyses were carried out using the Kaplan–Meier method and Cox proportional hazards regression models. Correlations were assessed using Spearman's rank correlation coefficient. A P-value < 0.05 was considered statistically significant. Statistical significance is indicated as follows: ns, not significant; *,P < 0.05, **,P < 0.01,***, P < 0.001.

2：Results

2.1 Differentially expressed genes shared by ImmuneScore and StromalScore are

predominantly enriched for immune-related functions

To delineate the precise alterations in the gene expression profiles of immune and stromal components within the TME, differential expression analysis was performed between samples stratified into high- and low-score groups based on ImmuneScore and StromalScore, respectively. Based on ImmuneScore, a total of 1,142 DEGs were identified, comprising 1,097 up-regulated and 45 down-regulated genes (Fig. 1A, C, D). Similarly, analysis based on StromalScore yielded 1,297 DEGs, including 1,247 up-regulated and 50 down-regulated genes (Fig. 1B–D). Intersection analysis using a Venn diagram revealed 482 commonly up-regulated genes and 12 commonly down-regulated genes shared between the ImmuneScore and StromalScore comparisons. These 494 overlapping DEGs were considered to be intimately associated with TME status. To further explore the underlying biological mechanisms, a PPI network was constructed using the STRING database and visualized with Cytoscape (National Institute of General Medical Sciences, NIGMS), which demonstrated extensive interactions among these 494 genes (Fig. 1E). GO enrichment analysis indicated that these DEGs were predominantly enriched in immune-related biological processes, such as leukocyte proliferation and B cell activation (Fig. 1F). KEGG enrichment analysis revealed that these genes were primarily involved in pathways including cytokine–cytokine receptor interaction, viral protein interaction with cytokine and cytokine receptor, and hematopoietic cell lineage (Fig. 1G). GSEA further demonstrated significant enrichment of these genes in immune-associated pathways, including innate immunity, leukocyte cell–cell adhesion, leukocyte-mediated immunity, regulation of lymphocyte activation, regulation of T cell activation, and T cell activation (Fig. 1H). Collectively, these findings suggest that the overall functional repertoire of TME-associated DEGs is predominantly linked to immune regulation, and that immune factors may constitute a defining feature of the TME in PCa.

Based on the 494 TME-associated genes described above, univariate Cox regression analysis was subsequently performed, identifying 32 genes significantly associated with BCR risk ($P < 0.05$). These genes were retained for the subsequent construction of the prognostic model (Fig.1I).

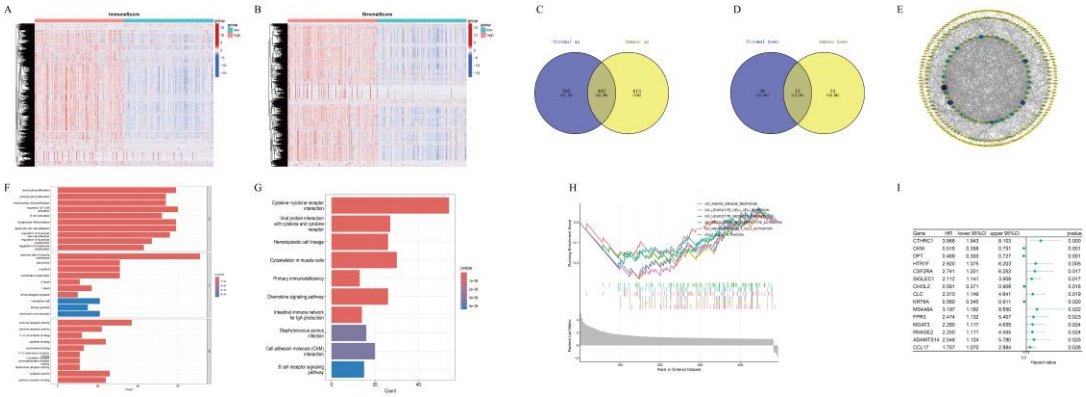
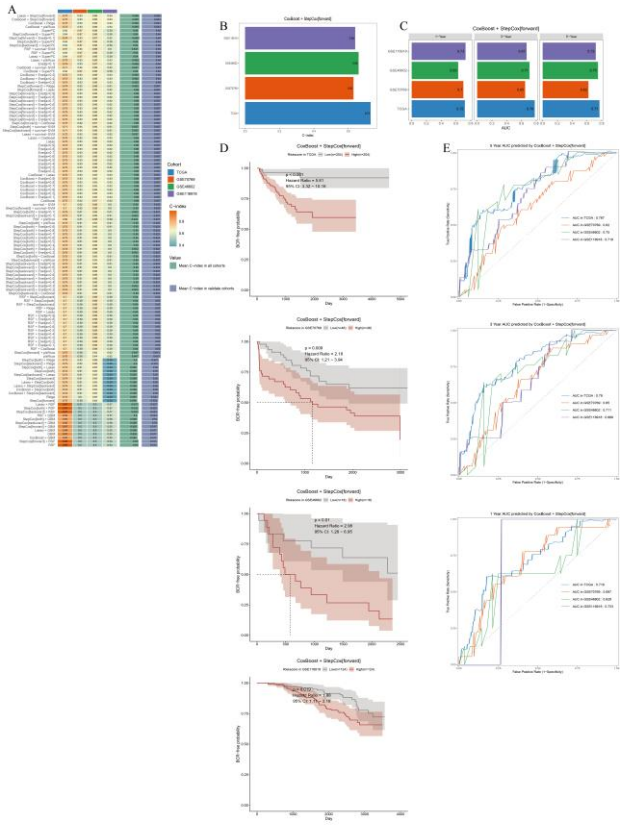


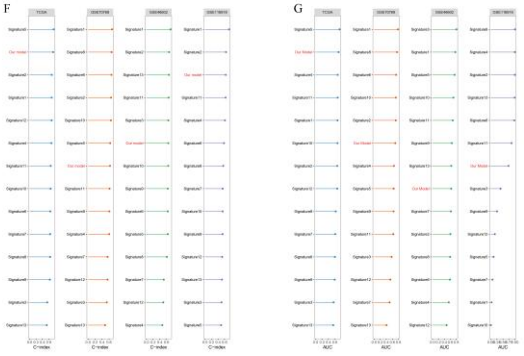
Fig 1. Identification and functional annotation of TME-related genes shared by the ImmuneScore and StromalScore groups in PCa.(A, B) Differential expression analyses between the high- and low-ImmuneScore groups and between the high- and low-StromalScore groups. DEGs were identified using the thresholds of adjusted $P < 0.05$ and $|\log2 \text{ fold change}| > 1$.(C, D) Venn diagrams showing the commonly upregulated and downregulated DEGs shared between the two comparisons, including 482 commonly upregulated genes and 12 commonly downregulated genes.(E) Protein-protein interaction (PPI) network of the 494 shared DEGs constructed using the STRING database and visualized with Cytoscape.(F, G) GO and KEGG enrichment analyses of the 494 shared DEGs.(H) GSEA showing that the TME-related genes were mainly enriched in immune-related pathways, including innate immune response, leukocyte adhesion, leukocyte-mediated immunity, regulation of lymphocyte activation, and T-cell activation.(I) Univariate Cox regression analysis of the 494 shared DEGs identified 32 candidate genes significantly associated with BCR risk ($P < 0.05$).

2.2 Construction and validation of the machine learning-based prognostic model

Based on the 32 genes significantly associated with BCR, a total of 17 combinations of machine learning algorithms were constructed and evaluated. The results demonstrated that Lasso + StepCox (forward), CoxBoost + StepCox (forward), CoxBoost + Ridge, and CoxBoost + Lasso were the four best-performing combinations. Ultimately, the CoxBoost + StepCox (forward) model was selected for subsequent analyses (Fig. 2A–B). This model comprised eight constituent genes: CCL17, CHI3L2, CSF2RA, CTHRC1, DPT, MNDA, SFRP2, and SPIB. The model exhibited stable

predictive performance for BCR outcomes at 1, 3, and 5 years across the three datasets (Fig. 2C). Kaplan–Meier analysis revealed a significantly elevated incidence of BCR among patients in the high-risk group (Fig. 2D), and ROC analysis further corroborated the predictive capacity of the model (Fig. 2E). Compared with 13 previously published models, the present model demonstrated superior performance in terms of C-index as well as 3- and 5-year AUC values; however, the 1-year AUC in the test set was comparatively lower (Fig. 2F–J).


A
B
C
D
E


F
G

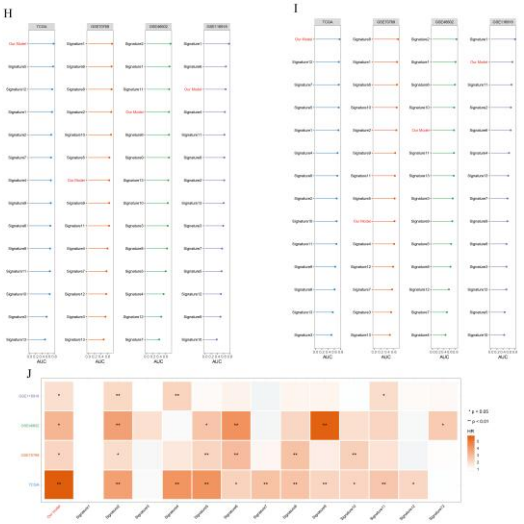
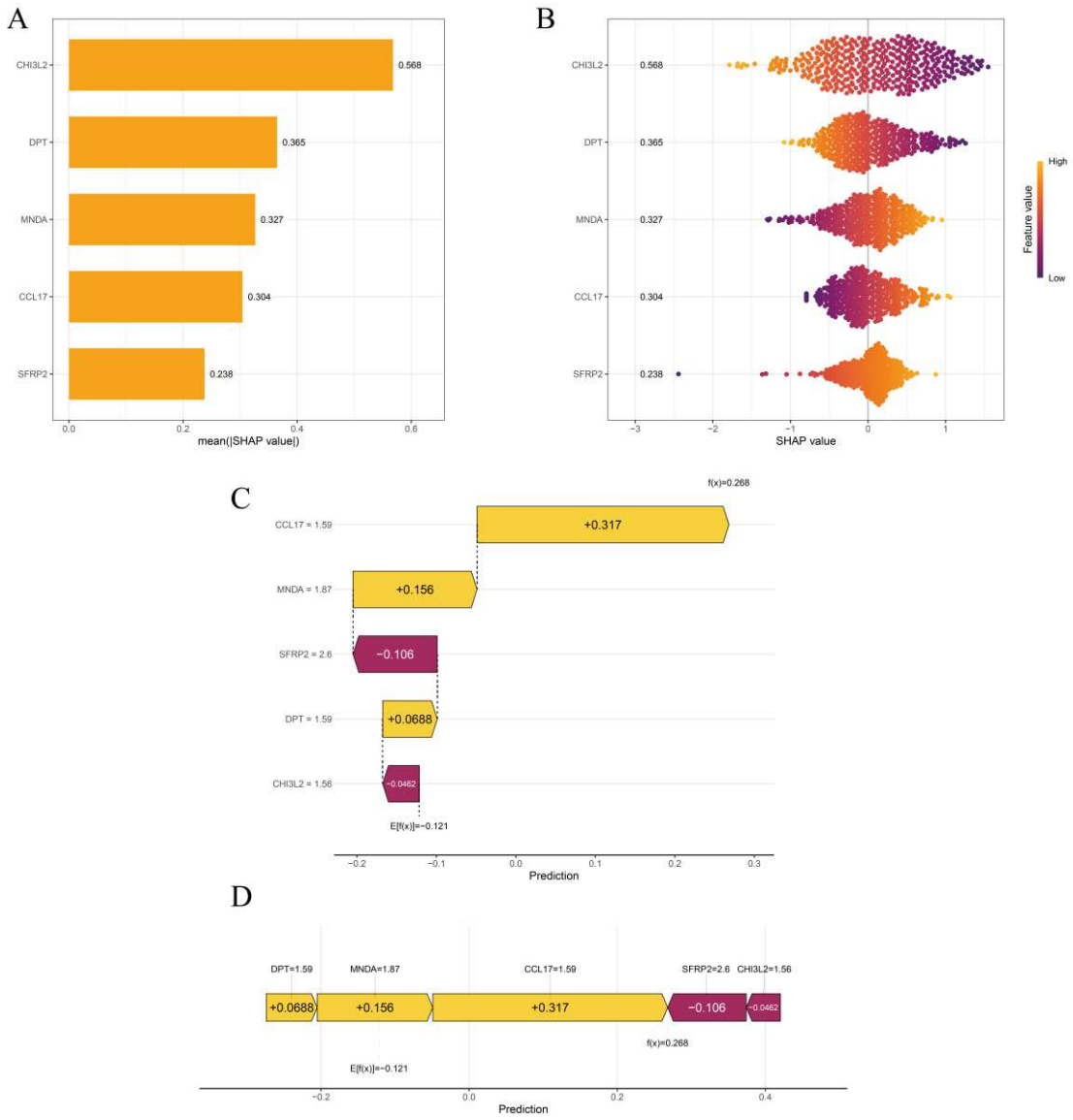

H
I
J

Fig 2. Construction and validation of the machine learning-based prognostic model Construction and validation of a machine learning-based prognostic model for BCR.(A) Heatmap showing the mean C-index values of 117 machine learning algorithm combinations developed from 32 genes significantly associated with BCR in the training and validation cohorts.(B) Comparison of the predictive performance of the four best-performing algorithm combinations, with CoxBoost + StepCox[forward] ultimately selected as the optimal model for subsequent analyses.(C) Time-dependent AUCs for predicting 1-, 3-, and 5-year BCR outcomes in three cohorts using the prognostic model composed of eight genes: CCL17, CHI3L2, CSF2RA, CTHRC1, DPT, MNDA, SFRP2, and SPIB.(D) Kaplan-Meier curves showing that patients in the high-risk group had a significantly higher probability of BCR than those in the low-risk group.(E) Time-dependent ROC curves for predicting 1-, 3-, and 5-year BCR, demonstrating the discriminative ability of the model.(F-J) Comparison of the predictive performance of the present model with 13 previously published models in terms of the C-index and AUCs at 1, 3, and 5 years.

2.3 SHAP-based model interpretability analysis

SHAP analysis was performed to interpret the contributions of the core genes. Based on mean absolute SHAP values, five core genes were identified: CHI3L2, DPT, MNDA, CCL17, and SFRP2 (Fig. 3A-B). SHAP summary and dependence plots further delineated the direction and magnitude of each gene's influence on risk prediction (Fig. 3C-D).



304

305

306 **Fig 3.**SHAP-based interpretability analysis of the prognostic model.(A, B) Ranking of model genes
307 according to the mean absolute SHAP values, identifying five core genes, including CHI3L2, DPT,
308 MNDA, CCL17, and SFRP2.(C, D) SHAP plots further illustrate the direction and magnitude of the
309 contribution of each core gene to risk prediction. Each dot represents one sample, the color indicates
310 gene expression level, and the SHAP value reflects the positive or negative contribution of the gene
311 to the model output.

2.4 Clinical characterization of the risk score

To evaluate the association between the prognostic risk score and clinicopathological features, the researchers conducted a correlation analysis. The results showed that patients who experienced BCR had significantly higher risk scores than those who did not (Fig. 4A). A higher risk score was significantly associated with several adverse clinicopathological characteristics, including advanced pathological T stage, advanced clinical T stage, advanced lymph node metastasis stage, advanced distant metastasis stage, higher Gleason score, PSA level, and age. (Fig. 4B–H)

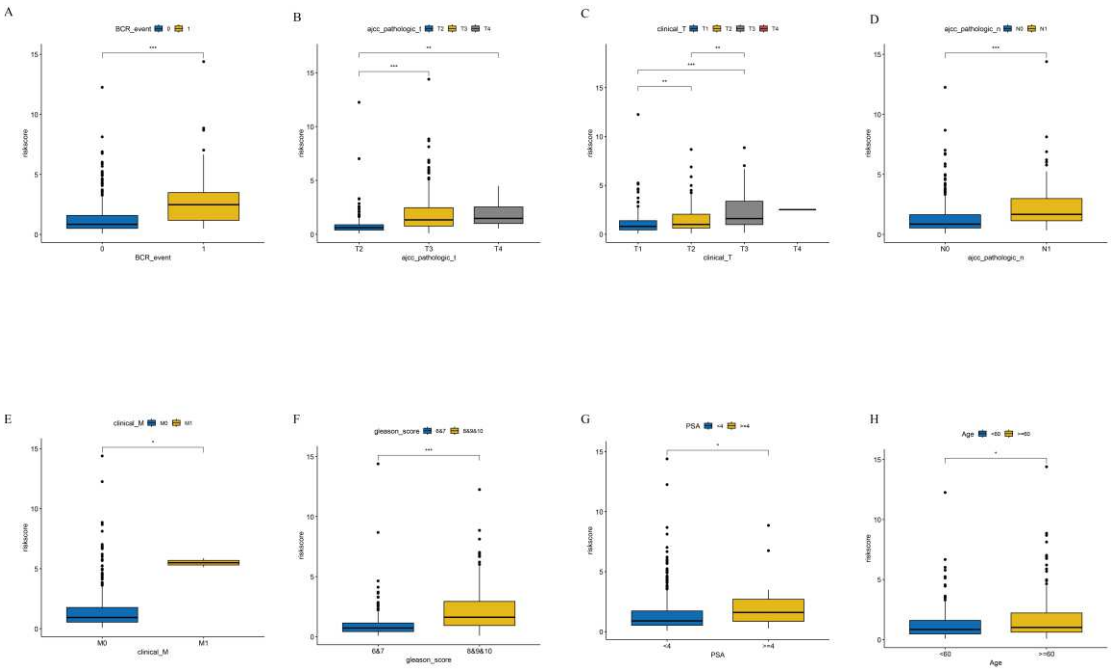


Fig 4. Correlation analysis of the risk score with clinicopathological characteristics.(A) Comparison of risk scores between patients with BCR and those without recurrence. Patients with BCR had significantly higher risk scores than non-recurrent patients.(B–H) Association of the risk score with various clinicopathological parameters, including pathological T stage (B), clinical T stage (C), lymph node metastasis (N) stage (D), distant metastasis (M) stage (E), Gleason score (F), PSA level (G), and age (H). A higher risk score was significantly associated with advanced pathological T stage, advanced clinical T stage, advanced N stage, advanced M stage, higher Gleason score, elevated PSA level, and older age.

2.5 Nomogram construction

Subsequently, univariate and multivariate Cox regression analyses were performed to identify independent predictors of BCR (Fig. 5A). Based on these results, a nomogram was constructed by integrating the risk score, Gleason score, pathological T stage, pathological N stage, and PSA level (Fig. 5B). Moreover, our risk score demonstrated a higher AUC value compared with traditional prognostic scoring systems (Fig. 5C). Calibration curves showed that the nomogram exhibited good predictive consistency (Fig. 5D–F).

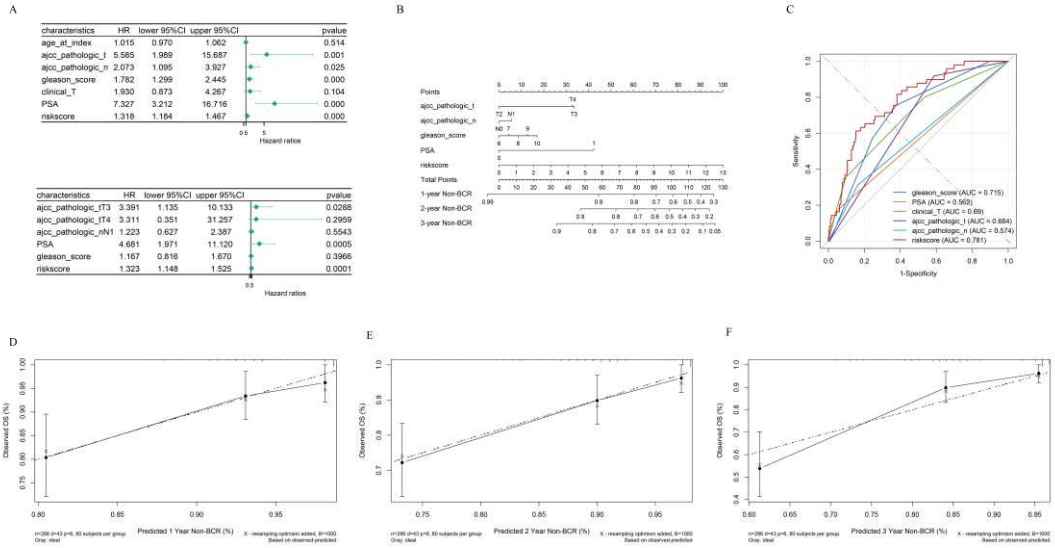


Fig 5. Identification of independent prognostic factors for BCR and construction of a nomogram.(A) Univariate and multivariate Cox regression analyses for identifying independent predictors of BCR.(B) Nomogram integrating the risk score, Gleason score, pathological T stage, pathological N stage, and PSA level for predicting the probability of BCR at 1, 3, and 5 years. (C) The risk score achieved a higher AUC value compared with traditional prognostic scoring systems. (D–F) Calibration curves of the nomogram for 1, 3, and 5-year BCR-free prediction, showing good consistency between the predicted and observed outcomes.

2.6 Comparison of immune infiltration between risk subgroups

GSEA was performed on the DEGs to elucidate the functional discrepancies between the risk subgroups. The analysis revealed numerous immune-related pathways within the MSigDB Collection (c5.cp.v7.0.symbols.gmt), including adaptive immune response, leukocyte differentiation, lymphocyte differentiation, regulation of lymphocyte activation, regulation of T cell activation, and T cell activation (Fig. 6A). To further characterize the immune microenvironmental features across risk subgroups, ESTIMATE, ssGSEA, CIBERSORT, and xCell analyses were conducted. ESTIMATE analysis demonstrated that the high-risk group exhibited significantly elevated stromal, immune, and ESTIMATE scores, whereas tumor purity was correspondingly lower compared with the low-risk group (Fig. 6B–E). ssGSEA results indicated that 25 immune cell subtypes—including activated CD4 T cells, central memory CD8 T cells, effector memory CD4 T cells, effector memory CD8 T cells, immature B cells, memory B cells, and regulatory T cells—were enriched at higher levels in the high-risk group (Fig. 6F). CIBERSORT analysis further revealed that the high-risk group harbored a greater proportion of M2 macrophages and a lower proportion of plasma cells and activated dendritic cells (Fig. 6G). xCell analysis yielded largely concordant findings, showing elevated proportions of chondrocytes, dendritic cells, hematopoietic stem cells, immature dendritic cells, macrophages, M1 macrophages, M2 macrophages, pericytes, and type 2 helper T cells in the high-risk group (Fig. 6H). Furthermore, correlation analysis confirmed significant associations between the expression of key model genes and the abundance of various immune cell types, thereby reinforcing the biological link between the BCR signature and the local immune milieu (Fig. 6I).

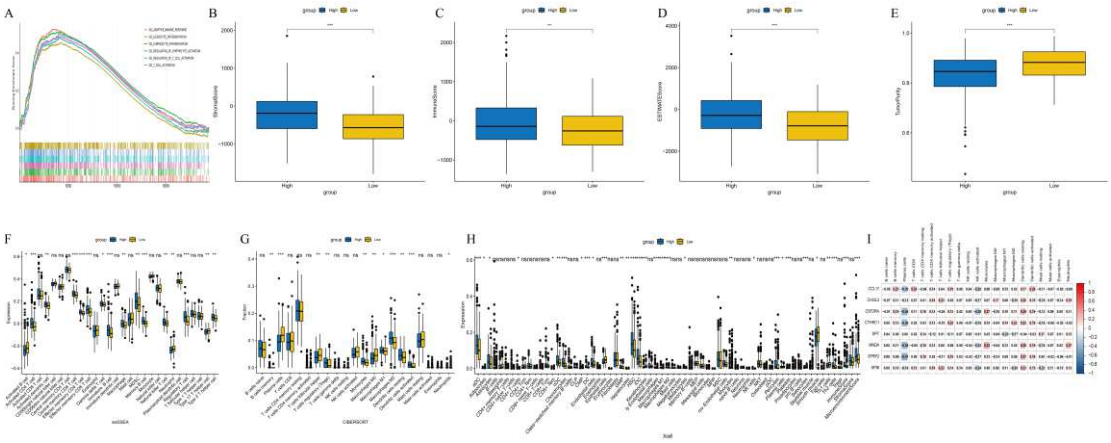


Fig 6. Functional enrichment and immune microenvironment characteristics in different risk subgroups.(A) GSEA showing that the differentially expressed genes between the high- and low-risk groups were mainly enriched in immune-related pathways, including adaptive immune response, leukocyte differentiation, lymphocyte differentiation, regulation of lymphocyte activation, regulation of T-cell activation, and T-cell activation.(B-E) ESTIMATE analysis comparing stromal score, immune score, ESTIMATE score, and tumor purity between the two risk groups. The high-risk group exhibited higher stromal, immune, and ESTIMATE scores, but lower tumor purity.(F) ssGSEA showing significant enrichment of multiple immune cell subtypes in the high-risk group, including activated CD4 T cells, central memory CD8 T cells, effector memory CD4 T cells, effector memory CD8 T cells, immature B cells, memory B cells, and regulatory T cells.(G) CIBERSORT analysis showing that the high-risk group had a higher proportion of M2 macrophages but lower proportions of plasma cells and activated dendritic cells.(H) xCell analysis further confirmed the differential immune infiltration patterns between the risk groups, with higher proportions of chondrocytes, dendritic cells, hematopoietic stem cells, immature dendritic cells, macrophages, M1 macrophages, M2 macrophages, pericytes, and type 2 helper T cells in the high-risk group.(I) Correlation analysis between key model genes and the abundance of multiple immune cell types.

2.7 Assessment of immunotherapeutic sensitivity and TIP analysis across risk subgroups

To evaluate potential differences in response to immunotherapy between risk subgroups, the expression levels of immune-regulatory targets were compared. The results demonstrated that the

majority of immune-regulatory molecules were significantly up-regulated in the high-risk group, including PDCD1LG2, CTLA4, CD80, CD86, TIM3, HAVCR2, TIGIT, TNFSF4, TNFSF9, ICOS, CD27, ICOSLG, TNFSF18, and CD70 (Fig. 7A–D), suggesting that the high- and low-risk groups may exhibit distinct patterns of response to immune checkpoint inhibitors. The TIP platform was further employed to evaluate the activity of the seven steps of the cancer-immunity cycle. The analysis revealed that the high-risk group displayed higher activity in Steps 1 and 4—namely, immune priming and immune cell recruitment—whereas no pronounced differences were observed in the subsequent effector phases (Fig. 7E). Collectively, these findings suggest that although the high-risk group exhibits more active immune recruitment and immunomodulatory signaling, anti-tumor immune effector functions may remain constrained, reflecting an immunologically inflamed yet functionally restricted microenvironment.

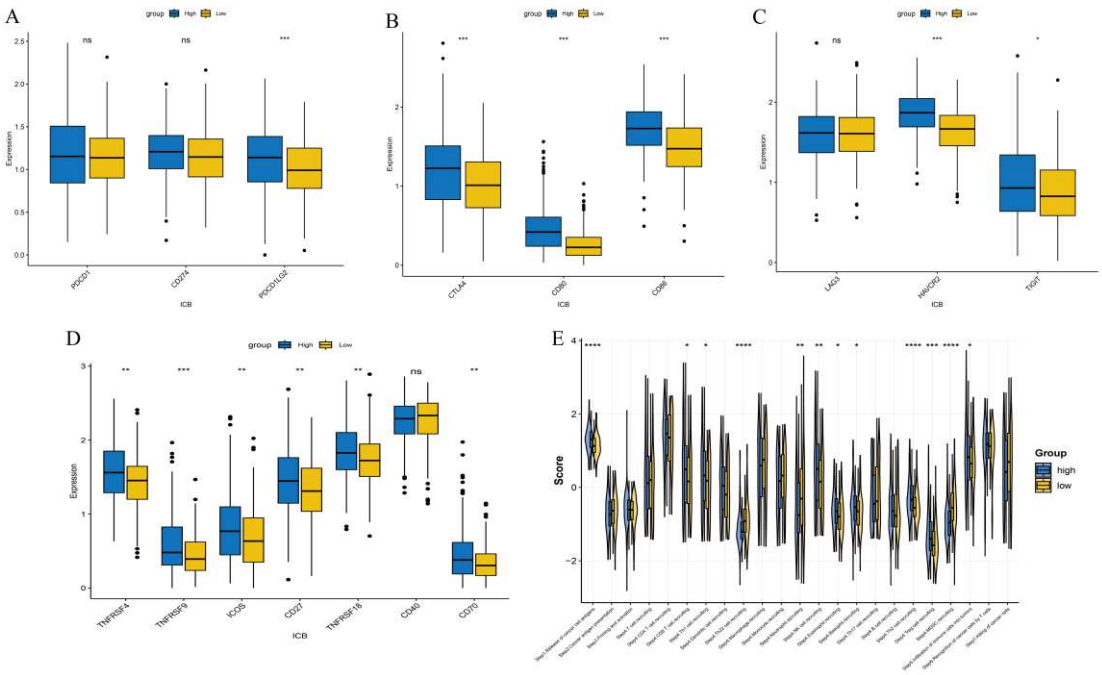


Fig 7. Comparison of immunomodulatory target expression and cancer-immunity cycle activity between different risk subgroups.(A-D) Differential expression of immunomodulatory molecules between the high-risk and low-risk groups, including PDCD1LG2, CTLA4, CD80, CD86, HAVCR2, TIGIT, TNFSF4, TNFSF9, ICOS, CD27, TNFSF18, and CD70.(E) Activity of the seven steps of

the cancer-immunity cycle evaluated using the TIP platform in the high-risk and low-risk groups.

2.8 Comparison of mutational landscapes between risk subgroups

Distinct gene mutations exert differential effects on the efficacy of immunotherapy; therefore, the mutational landscape of PCa was assessed. The mutational profiles of the risk subgroups are illustrated in Fig. 8A and 8B. TP53 mutations predominated in the high-risk group, whereas TTN mutations were more prevalent in the low-risk group.

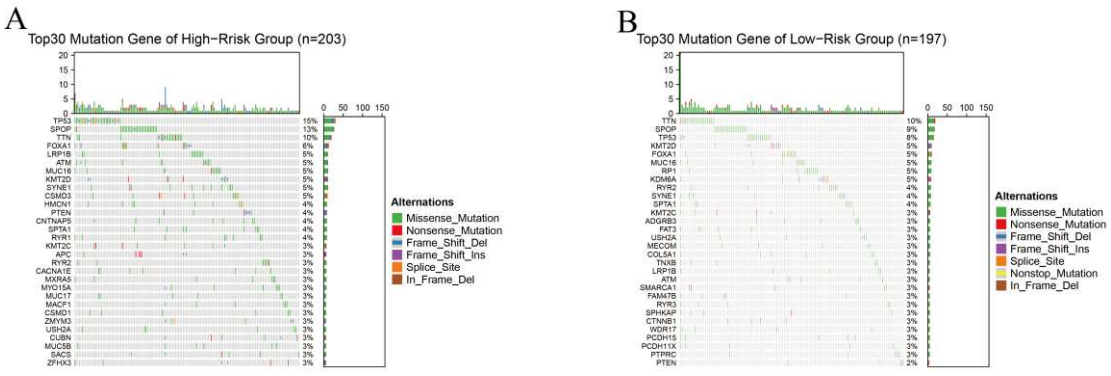


Fig 8. Comparison of the somatic mutation landscape between different risk subgroups.(A, B) Somatic mutation profiles of the high-risk and low-risk groups. Waterfall plots show the distribution patterns and mutation frequencies of mutated genes in each subgroup.

2.9 Identification of hub genes associated with risk subgroups and immunity via WGCNA

A total of 441 DEGs were identified between the two risk subgroups, comprising 263 up-regulated and 178 down-regulated genes (Fig. 9A), and these genes were utilized for WGCNA (Fig. 9B–C). Module–trait correlation analysis revealed that the blue module exhibited significant

correlations with both risk subgroup ($R = 0.59$, $P = 3 \times 10^{-40}$) and immune characteristics ($R = 0.47$, $P = 4 \times 10^{-24}$), and was therefore designated as the key module (Fig. 9D). Within this module, 16 hub genes were identified based on the criteria of MM > 0.7 and GS > 0.35, including INHBA, CTHRC1, NOX4, ADAMTS2, ITGA11, COL1A1, FAP, POSTN, COL8A1, HTR2B, BGN, COL5A2, COL3A1, VCAN, TMEM130, and FAM180A (Fig. 9E).

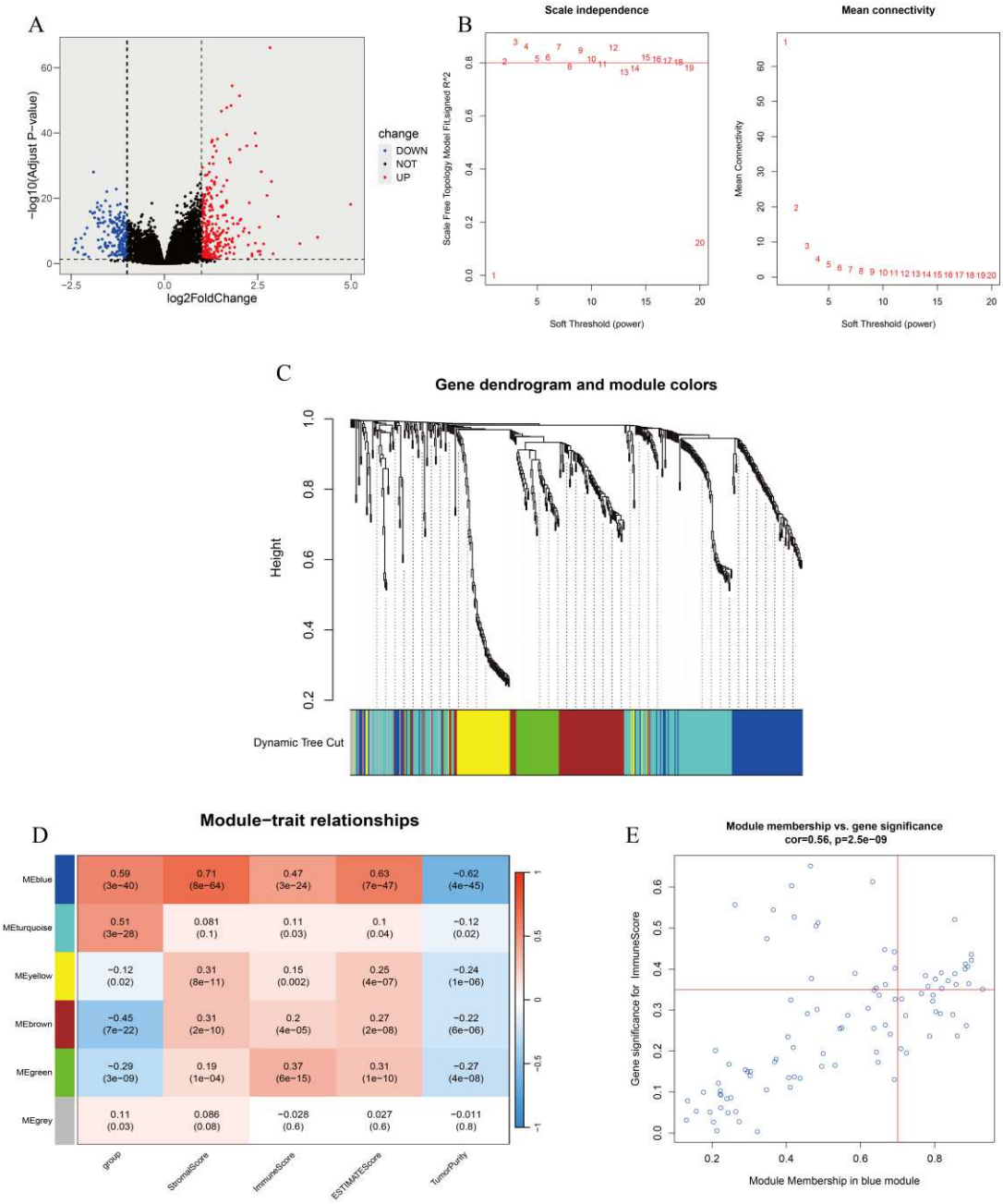


Fig 9. Identification of key modules and hub genes associated with risk subgroups and immune

features by WGCNA.(A) Volcano plot of differentially expressed genes between the high-risk and low-risk groups.(B) Soft-thresholding power analysis for construction of the scale-free co-expression network in WGCNA.(C)Gene dendrogram and module color assignment.(D) Heatmap of module-trait relationships showing the correlations of each module with the risk subgroup and immune features, with the blue module exhibiting significant associations with both traits.(E) Scatter plot of module membership (MM) versus gene significance (GS) in the blue module for identification of hub genes.

2.10 Functional enrichment of hub genes and their correlation with immune infiltration

To elucidate the potential functions of the aforementioned hub genes, GO enrichment analysis was performed. The results demonstrated that these genes were predominantly enriched in processes such as extracellular matrix organization, collagen fibril organization, and collagen metabolism (Fig. 10A). These findings suggest that the immunological differences between risk subgroups may not be directly driven by immune-related genes per se but are more likely mediated indirectly through stromal remodeling, collagen metabolism, and extracellular structural reorganization, which in turn influence immune cell infiltration and activation. PPI network and gene correlation analyses further revealed close interactions among these hub genes (Fig. 10B). Spearman correlation analysis demonstrated that the majority of hub genes were significantly associated with immune infiltration characteristics (Fig. 10C–E).

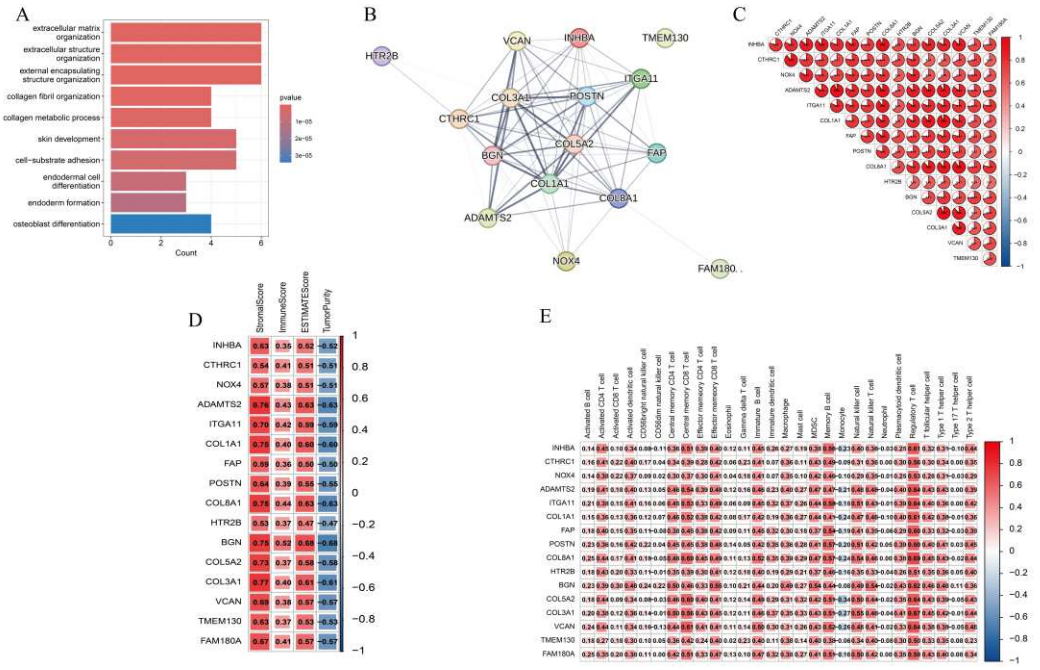


Fig 10. Functional enrichment, interaction network, and immune correlation analyses of hub genes.(A) Gene Ontology (GO) enrichment analysis of the hub genes.(B) Protein-protein interaction (PPI) network of the hub genes.(C) Spearman correlation analysis among the hub genes.(D) Correlation analysis between the hub genes and ESTIMATE-related features. (E) Correlation analysis between the hub genes and immune cell infiltration characteristics.

2.11 Comparison of drug sensitivity and TIDE scores between risk subgroups

In silico drug sensitivity analysis was performed to identify compounds with significantly differential efficacy between risk subgroups. The results indicated that patients in the high-risk group exhibited significantly greater sensitivity to commonly used chemotherapeutic agents such as oxaliplatin and cyclophosphamide (Fig. 11A–B). To evaluate immunotherapeutic response characteristics, TIDE scores were calculated. The high-risk group exhibited significantly elevated TIDE scores (Fig. 11C), suggesting a greater propensity for immune evasion and potential resistance to immune checkpoint inhibitors.

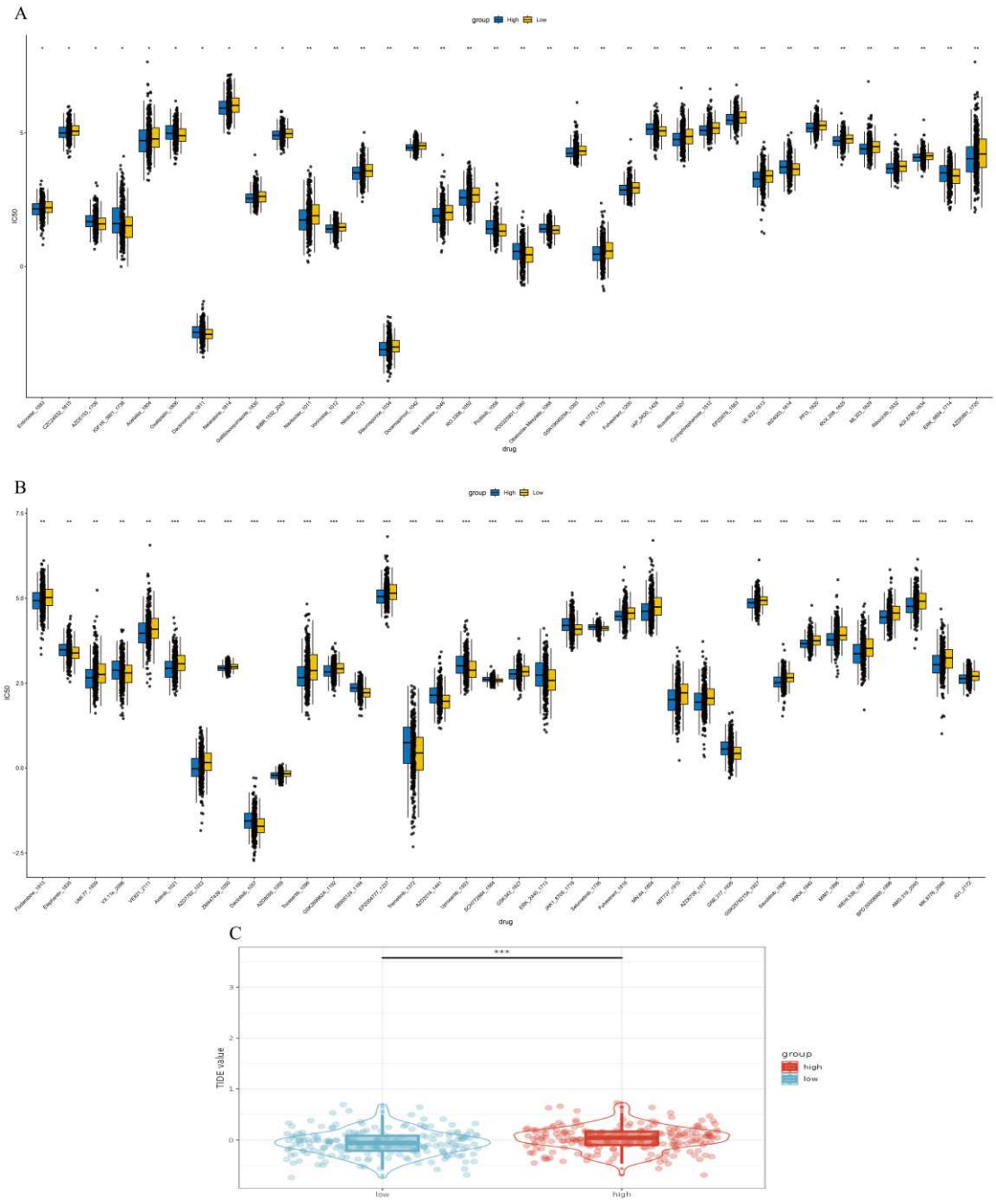


Fig 11. Drug sensitivity and immunotherapy response characteristics in different risk subgroups.(A, B) Comparison of the predicted sensitivity to chemotherapeutic agents between the high-risk and low-risk groups.(C)Comparison of TIDE scores between the two groups for evaluating potential immune evasion and response to immune checkpoint blockade therapy.Statistical significance is indicated as follows: ns, not significant; $P < 0.05$, $P < 0.01$, $P < 0.001$.

2.12 Annotation of cell subpopulations and identification of predominant subpopulations

expressing risk genes

To further enhance the biological interpretability of the model, single-cell RNA sequencing data of PCa were normalized and subjected to dimensionality reduction using uniform manifold approximation and projection (UMAP). A total of 5,822 cells were clustered (Fig. 12A–B). Cell type annotation was performed using the scMayoMap function, which identified seven major cell types: common myeloid progenitors, endothelial cells, epithelial cells, macrophages, smooth muscle cells, T lymphocytes, and tissue stem cells (Fig. 12C). Differential expression analysis further elucidated the expression patterns of DPT, MNDA, SFRP2, CSF2RA, and CTHRC1 across these cell types. The results revealed that DPT, SFRP2, and CTHRC1 were highly expressed in epithelial cells, whereas MNDA and CSF2RA were predominantly expressed in macrophages (Fig. 12D–E). These observations suggest that distinct risk genes may participate in the regulation of the PCa microenvironment through different cellular components.

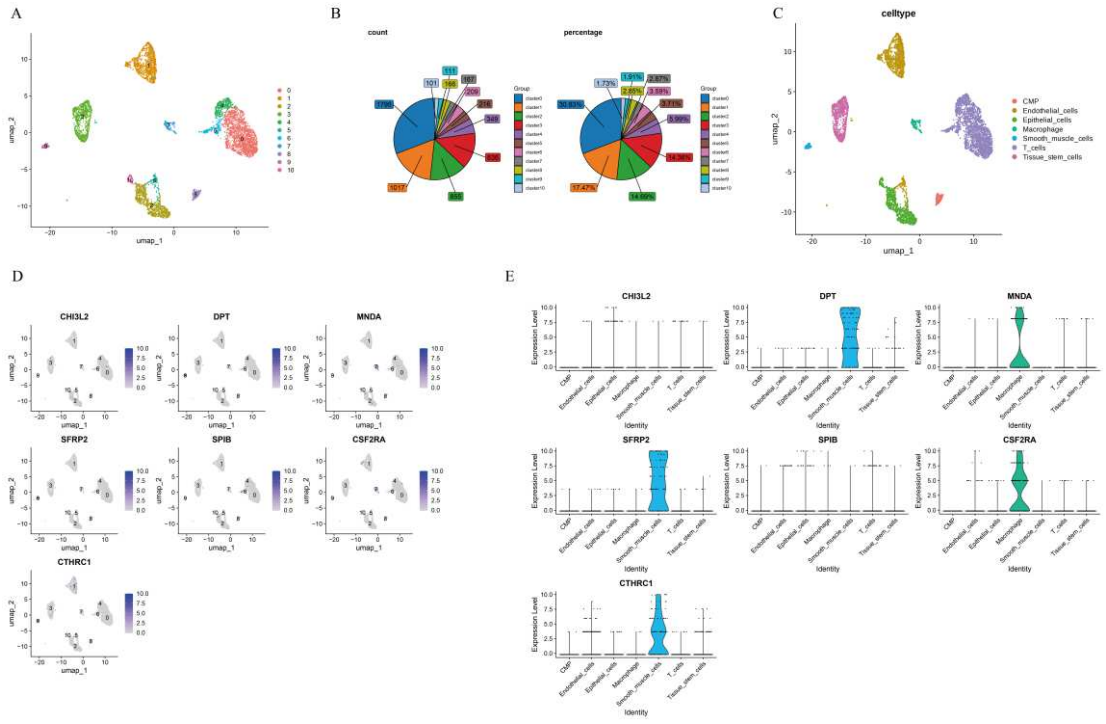


Fig 12. Single-cell transcriptomic analysis reveals the expression patterns of prognostic genes across different cell types.(A, B)UMAP visualization of the PCa single-cell transcriptomic dataset after

normalization, dimensionality reduction, and clustering.(C) Cell type annotation using scMayoMap identified seven major cell types, including common myeloid progenitors, endothelial cells, epithelial cells, macrophages, smooth muscle cells, T lymphocytes, and tissue stem cells.(D, E) Expression distributions of key model genes across different cell types. DPT, SFRP2, and CTHRC1 were mainly expressed in epithelial cells, whereas MNDA and CSF2RA were predominantly expressed in macrophages.

2.13 Pathway analysis of risk genes at the single-cell level

To delineate the potential pathogenic pathways in which the risk genes may participate within their predominant subpopulations, pathway activity profiling was performed using scGSVA. Hallmark pathway enrichment analysis revealed that CHI3L2 in tissue stem cells was primarily enriched in the apical junction pathway, which is closely associated with cell adhesion, polarity maintenance, and stemness regulation (Fig. 13A). In smooth muscle cells, DPT, SPIB, SFRP2, and CTHRC1 exhibited significant pathway enrichment, participating in the remodeling of the PCa stromal microenvironment through pathways related to hormone response, metabolic reprogramming, inflammatory regulation, cell adhesion, and invasion. Specifically, DPT was enriched in multiple pathways, including early estrogen response, coagulation, reactive oxygen species, Wnt/β-catenin signaling, glycolysis, and fatty acid metabolism. SPIB was primarily enriched in heme metabolism and p53 signaling. SFRP2 was enriched in pathways associated with inflammatory immune response, angiogenesis, epithelial-mesenchymal transition, and hormone and lipid metabolism. CTHRC1 was primarily enriched in apical junction, epithelial-mesenchymal transition, glycolysis, IL6-JAK-STAT3 signaling, and TGF-β signaling pathways (Fig. 13B, 13D, 13E, 13G). In macrophages, MNDA exhibited significant positive correlations with the interferon-α response pathway and pancreatic beta cell-related pathways, whereas CSF2RA was significantly negatively correlated with the coagulation pathway and late estrogen response pathway. Both genes are involved in the functional regulation of tumor-associated macrophages and the maintenance of immune microenvironmental homeostasis in PCa (Fig. 13C, 13F).

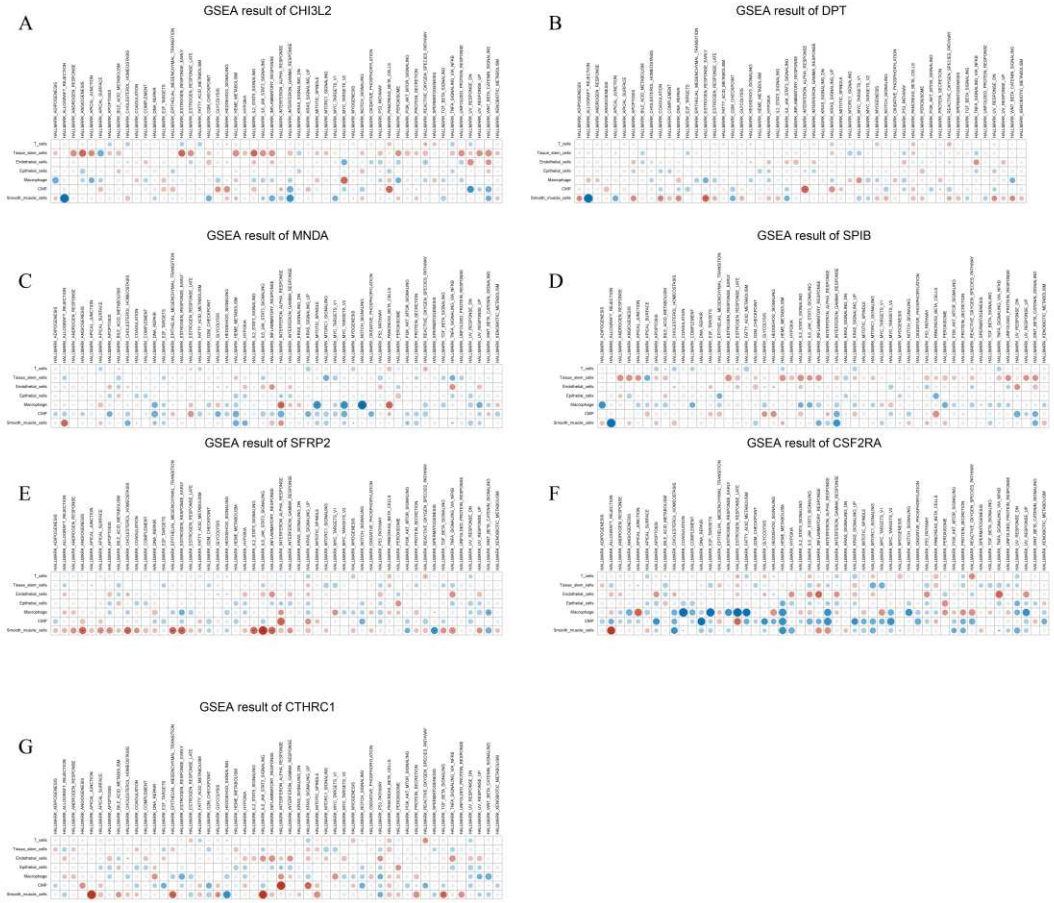


Fig 13.scGSVA analysis of pathway activities of prognostic genes in dominant cell subpopulations.(A) CHI3L2 in tissue stem cells was mainly enriched in the apical junction pathway.(B, D, E, G) In smooth muscle cells, DPT, SPIB, SFRP2, and CTHRC1 were enriched in pathways related to hormone response, metabolic reprogramming, inflammatory regulation, cell adhesion, and invasion.(C, F) In macrophages, MNDA was positively associated with the interferon-α response and pancreas beta cells-related pathways, whereas CSF2RA was negatively associated with coagulation and late estrogen response pathways.

2.14 In silico knockout analysis of MNDA and DPT

To further resolve the roles of MNDA and DPT—two risk genes whose functions in BCR remain poorly characterized—in silico knockout analyses were performed. Using an adjusted P-value < 0.05 as the threshold, 8 and 16 stable knockout-responsive genes were identified for MNDA and DPT, respectively (Fig. 14A–B, 14D–E). GO enrichment analysis revealed that the knockout-responsive genes of MNDA were significantly enriched in extracellular matrix assembly, extracellular structure organization, and external encapsulating structure organization (Fig. 14C). The knockout-responsive genes of DPT were significantly enriched in lipoprotein metabolic processes (Fig. 14F). These genes may represent key regulatory factors through which MNDA and DPT mediate BCR, TME remodeling, and malignant progression in PCa.

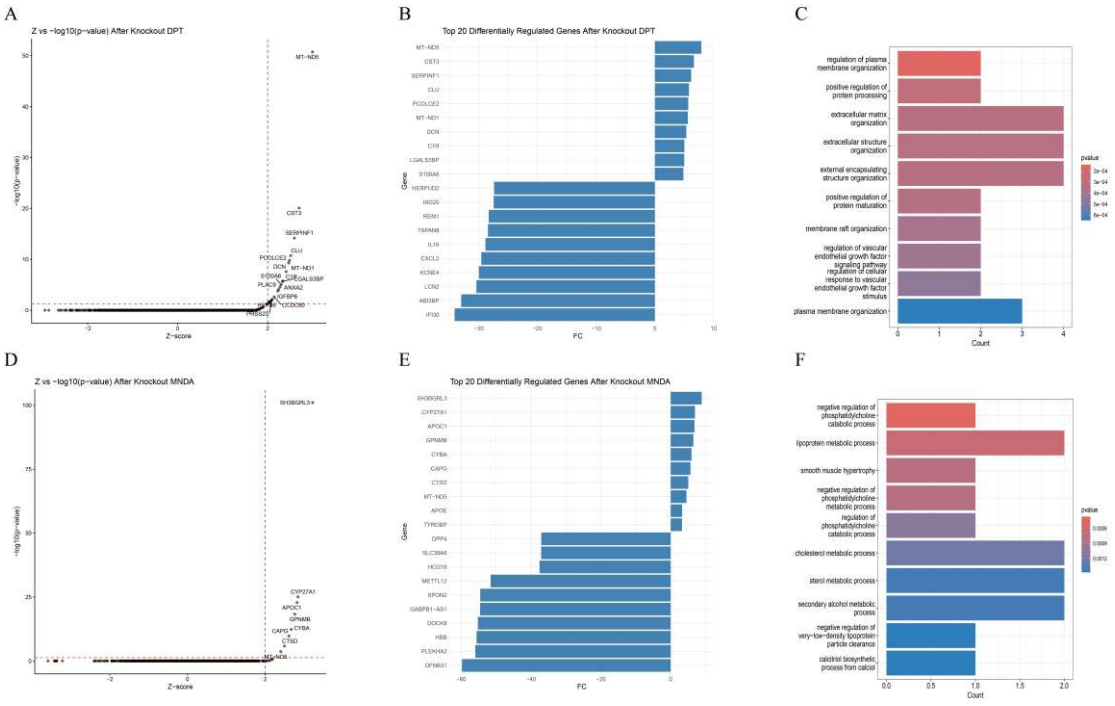


Fig 14. In silico knockout simulation and functional enrichment analyses of MNDA and DPT. (A, B) Stable KO-responsive genes identified after in silico knockout of MNDA.(C)GO enrichment analysis of the MNDA KO-responsive genes, showing significant enrichment in extracellular matrix assembly, extracellular structure organization, and external encapsulating structure organization.(D, E) Stable KO-responsive genes identified after in silico knockout of DPT.(F) GO enrichment analysis of the DPT KO-responsive genes, showing significant enrichment in lipoprotein metabolic processes.These findings suggest that MNDA and DPT may contribute to BCR, tumor microenvironment remodeling, and malignant progression in PCa through regulation of

extracellular matrix remodeling- or metabolism-related pathways.

3. Discussion

Substantial evidence has accrued indicating that BCR exerts a profound impact on the long-term prognosis of patients with PCa[25]. Once BCR occurs, radiotherapy represents the principal therapeutic option[26]. Studies have demonstrated that the initiation of salvage radiotherapy at a prostate-specific antigen (PSA) level below 0.5 ng/mL is significantly associated with prolonged metastasis-free survival[27]. However, the paucity of reliable biomarkers may result in the overtreatment of patients with radiotherapy, thereby increasing the risk of unnecessary adverse effects[28]. Follow-up imaging frequently fails to detect recurrent disease, and serial serum PSA monitoring remains the primary modality for detecting BCR following primary treatment[29]. This underscores an unmet clinical need for biomarkers capable of refining risk stratification for BCR in patients with PCa. Furthermore, the role of the TME in PCa initiation, progression, immune evasion, and therapeutic response has garnered widespread attention. Consequently, a detailed characterization of BCR risk from the perspective of TME-associated genes is of critical importance.

In the present study, a multi-algorithm machine learning integration strategy was employed to construct and validate an eight-gene prognostic model for predicting BCR following RP in patients with PCa, based on TME-associated genes. The eight genes comprising the model are CCL17, CHI3L2, CSF2RA, CTHRC1, DPT, MNDA, SFRP2, and SPIB. Among these, SPIB, CHI3L2, DPT, and CSF2RA were negatively correlated with BCR risk, whereas the expression of CCL17, CTHRC1, MNDA, and SFRP2 exhibited positive correlations with BCR risk. Notably, DPT was expressed at lower levels in patients who experienced BCR. The model demonstrated robust discriminatory capacity in both the training cohort and external validation datasets, achieving a C-index superior to that of multiple previously published prognostic models for PCa. The risk score was significantly associated with adverse clinicopathological features, including elevated Gleason score, lymph node metastasis, and distant metastasis, suggesting that the model captures not merely

recurrence risk but rather a more aggressive tumor biological state with a pronounced propensity for progression.

A noteworthy paradoxical phenomenon was observed: although the high-risk group exhibited elevated immune scores and more abundant immune cell infiltration, this was accompanied by a poorer prognosis. Specifically, subpopulations including activated $CD4^+$ T cells, central memory $CD8^+$ T cells, effector memory T cells, B cells, dendritic cells, and macrophages were significantly increased in the high-risk group. This phenomenon is not unique to PCa and has been similarly documented in various other solid tumors, including breast cancer and lung cancer [30-31].

Within the high-risk group, the proportions of M2 macrophages and regulatory T cells (Tregs) were significantly elevated, whereas the proportions of plasma cells and activated dendritic cells were diminished. CIBERSORT analysis revealed a higher abundance of M2 macrophages in the high-risk group. Previous studies have indicated that M2 macrophages in PCa are generally associated with immunosuppression, tissue remodeling, and adverse prognosis[32-33]. The reduction of activated dendritic cells, which serve as key antigen-presenting cells, may impair the initiation and maintenance of effective T cell responses. Concurrently, ssGSEA indicated a relative increase in Tregs within the high-risk group, and xCell analysis further corroborated an expansion of Th2 cells and various myeloid cell components. Collectively, these observations suggest that although the high-risk group is enriched for immune cells, this infiltration does not equate to effective anti-tumor immunity; rather, it may reflect a paradoxical state characterized by the coexistence of immune activation and immunosuppression.

TIP analysis demonstrated that the high-risk group exhibited greater activity in the early steps of the cancer-immunity cycle, particularly immune cell recruitment, whereas subsequent effector phases showed limited enhancement. This pattern implies that the immune response may be functionally impeded following recruitment or infiltration. TIDE analysis further revealed that the high-risk group had significantly elevated TIDE scores, indicative of more pronounced immune

dysfunction, a greater propensity for immune exclusion, and a potential risk of resistance to immune checkpoint inhibitors (ICIs). Taken together, the integrated findings from immune checkpoint expression, TIP, and TIDE analyses indicate that while the high-risk group manifests stronger immune-related signals and richer immune infiltration, this inflammatory state likely represents a TME phenotype accompanied by concurrent immunosuppression and functional impairment.

Through WGCNA, a module of stromal remodeling genes significantly associated with the risk subgroups was identified. This module predominantly comprised genes such as INHBA, CTHRC1, NOX4, ADAMTS2, ITGA11, COL1A1, FAP, POSTN, COL8A1, BGN, COL5A2, COL3A1, and VCAN. The majority of these genes are intimately linked to extracellular matrix (ECM) remodeling, collagen deposition, fibroblast activation, and stromal reaction. GO enrichment analysis further confirmed that these genes were significantly enriched in processes including ECM organization, collagen fibril organization, and collagen metabolism. Evidence suggests that excessive ECM deposition and stromal remodeling not only facilitate tumor cell migration and invasion but also restrict effective immune cell infiltration and effector function through physical barriers and altered biochemical signaling, thereby fostering immune exclusion or a state of functional immunosuppression[34-36]. Hence, stromal remodeling within the high-risk group may constitute a key mechanistic link between the high immune infiltration phenotype and the adverse clinical outcomes observed.

Analysis of the cellular origins of the model genes revealed that DPT, SFRP2, and CTHRC1 were predominantly expressed in epithelial cells, whereas MNDA and CSF2RA were primarily expressed in macrophages. These findings, derived from single-cell transcriptomic data, support the notion that the risk model reflects not merely intrinsic alterations within tumor cells but rather a broader process of microenvironmental reprogramming involving the coordinated participation of tumor cells, immune components, and stromal elements. Further scGSVA indicated that distinct key genes were enriched for different functional pathways within specific cell populations: CHI3L2 in tissue stem cells was associated with apical junction-related pathways; DPT, SFRP2, and CTHRC1 in smooth muscle cells were linked to pathways involving hormone response, inflammatory

regulation, glycolysis, epithelial-mesenchymal transition, and TGF-β/IL6-JAK-STAT3 signaling; and MNDA and CSF2RA in macrophages were correlated with interferon response, coagulation, and metabolism-related signals. In silico knockout analyses suggested that MNDA knockout-responsive genes were predominantly enriched in processes such as ECM assembly and extracellular structure organization, whereas DPT knockout-responsive genes were primarily enriched in lipoprotein metabolic processes. These findings provide clear directions for subsequent functional investigations, although further validation in both in vitro and in vivo models is warranted.

From a therapeutic perspective, the high-risk group exhibited higher predicted sensitivity to agents such as oxaliplatin and cyclophosphamide, and the expression of the majority of immune-regulatory targets—including PDCD1LG2, CTLA4, CD80, CD86, HAVCR2, TIGIT, ICOS, CD27, TNFSF4, TNFSF9, TNFSF18, and CD70—was significantly elevated. The upregulation of immune checkpoints may, on the one hand, indicate a more active background immune response in the high-risk group; on the other hand, it may reflect T cell exhaustion, enhanced immune negative feedback, and the emergence of tumor immune evasion under persistent antigen stimulation. However, the elevated TIDE scores suggest that ICI monotherapy may yield limited efficacy. Particularly for PCa, which is generally considered to exhibit low immunogenicity, reliance solely on elevated immune checkpoint molecule expression to infer ICI benefit remains subject to considerable limitations[37]. Given the prominent stromal remodeling features of the high-risk group and the advances in the application of stromal modifying agents in the treatment of solid tumors such as pancreatic and lung cancers, future strategies may warrant exploration of combination therapies integrating chemotherapy with FAK inhibitors, TGF-β inhibitors, and ICIs[38-39]. It is important to emphasize that these predictions of drug sensitivity and immunotherapeutic response are computationally derived and require validation in prospective clinical cohorts.

The precise functions of MNDA and DPT among the model genes in the context of BCR have yet to be fully elucidated. Myeloid nuclear differentiation antigen (MNDA) belongs to the human PYHIN (pyrin and hematopoietic interferon-inducible nuclear domain) protein family and is implicated in mediating inflammatory responses, cellular senescence, and the recognition of foreign

double-stranded DNA[40-42]. A study in hepatocellular carcinoma (HCC) indicated that MNDA may accelerate HCC metastasis by promoting M2 macrophage polarization and the secretion of pro-metastatic proteins. Within the immune microenvironment of colorectal cancer, MNDA has also been shown to mediate H3K18 lactylation via the MNDA/EP300-CXCR2 regulatory axis, thereby facilitating the infiltration of polymorphonuclear myeloid-derived suppressor cells and the establishment of an immunosuppressive microenvironment. Additionally, MNDA has been proposed as a potential diagnostic marker for marginal zone lymphoma[43]. Collectively, this evidence suggests that MNDA may contribute to immune evasion within the PCa TME. Dermatopontin (DPT), a non-collagenous ECM protein initially identified in dermal tissue, functions in regulating ECM structure and modulates the activity of decorin and transforming growth factor-beta (TGF-β)[44]. Studies in breast cancer, colorectal cancer, and cholangiocarcinoma have demonstrated that DPT may influence the immunosuppressive state of the TME by regulating ECM remodeling, mediating immune cell infiltration, and modulating chemokine secretion, and low DPT expression has been significantly associated with poor patient prognosis[45-47].

The present study has several limitations. First, the data were derived from TCGA and public databases, constituting a retrospective analysis that may be subject to inherent selection bias. Second, the biological functions of the model genes and their causal relationship with BCR lack experimental validation in vitro and in vivo. Although single-cell transcriptomic data were incorporated to provide complementary analysis of cell-type-specific expression of the model genes, not all model genes could be robustly identified at the single-cell level due to limitations in sequencing depth, capture efficiency of low-abundance transcripts, and cell annotation accuracy. Third, assessments pertaining to immunotherapy were based on indirect evidence from TIDE and TIP analyses, lacking clinical response data from real-world PCa immunotherapy cohorts. Fourth, drug sensitivity predictions were inferred from transcriptomic profiling and do not equate to actual therapeutic efficacy. Future investigations should integrate prospective cohorts, spatial transcriptomics, and functional experiments to systematically validate these findings.

In conclusion, the present study provides a TME-based risk stratification tool for predicting

BCR following RP in patients with PCa. It reveals a paradoxical phenotype in the high-risk group characterized by immune inflammation coupled with functional restriction, and proposes stromal remodeling as a key mechanistic explanation for this phenomenon. This model holds promise as a novel theoretical foundation for individualized prognostic assessment and the optimization of therapeutic strategies in patients with PCa.

Data availability statement

The transcriptome datasets analysed during the current study are derived from two publicly accessible repositories: The Cancer Genome Atlas (TCGA) and the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO). The specific datasets and their corresponding direct access links are as follows: TCGA-PRAD: Prostate Adenocarcinoma project of TCGA, available at https://portal.gdc.cancer.gov/projects/TCGA-PRAD. GSE70769: available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE70769. GSE46602: available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE46602. GSE116918: available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE116918. GSE157703: available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE157703.

All data from these repositories have been fully released and are publicly accessible for verification.

Ethics, Consent to Participate, and Consent to Publish declarations

not applicable

Authors' contributions

Taize Sun was responsible for data management, statistical analysis, and drafting of the initial manuscript. JunCheng Chen contributed to manuscript revision and critical review. Fei Wang provided overall supervision and guidance. All authors have read and approved the final version of the manuscript.

Funding

This research was supported by the National Natural Science Foundation of China (Grant No. 82460553) and the Academic Enhancement Support Program of Hainan Medical University (Grant No. XSTS2025086).

Competing interests

All authors declare that they have no competing interests.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

[1].BRAY F, LAVERSANNE M, SUNG H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. CA Cancer J Clin, 2024, 74(3): 229-63.

[2].ZHANG X, LI X, YU Q, et al. Morbidity and All-Cause Mortality Following Radical Prostatectomy Compared with Observation for Localized Prostate Cancer in Chinese Men: A Non-Randomized Retrospective Study [J]. Med Sci Monit, 2021, 27: e928596.

[3].TANG D G. Understanding and targeting prostate cancer cell heterogeneity and plasticity [J]. Semin Cancer Biol, 2022, 82: 68-93.

[4].PREISSER F, ABRAMS-POMPE R S, STELWAGEN P J, et al. European Association of Urology Biochemical Recurrence Risk Classification as a Decision Tool for Salvage Radiotherapy-A Multicenter Study [J]. Eur Urol, 2024, 85(2): 164-70.

[5].SUZUKI A, SATO S, NAKAIGAWA N, et al. Combination of Blood Adiponectin and Leptin Levels Is a Predictor of Biochemical Recurrence in Prostate Cancer Invading the Surrounding Adipose Tissue [J]. Int J Mol Sci, 2024, 25(16).

[6].ROY S, MORGAN S C. Who Dies From Prostate Cancer? An Analysis of the Surveillance, Epidemiology and End Results Database [J]. Clin Oncol (R Coll Radiol), 2019, 31(9): 630-6.

[7].LIU Y N, CHEN W Y, LIU M K, et al. Immunosuppressive role of BDNF in therapy‐induced neuroendocrine prostate cancer [J]. Molecular Oncology, 2024, 18(6): 1665-86.

[8].OWEN J S, CLAYTON A, PEARSON H B. Cancer-Associated Fibroblast Heterogeneity, Activation and Function: Implications for Prostate Cancer [J]. Biomolecules, 2022, 13(1).

[9].JI S, WU W, JIANG Q. Crosstalk between Endothelial Cells and Tumor Cells: A New Era in Prostate Cancer Progression [J]. Int J Mol Sci, 2023, 24(23).

[10].MARTINEZ-ROMERO A, MARTINEZ-LARRINAGA A, GREGO-BESSA J, et al. Aggressive prostate cancer is associated with pericyte dysfunction [J]. Mol Oncol, 2025.

[11].SATO A, TAKAGI K, YAMAGUCHI-TANAKA M, et al. Matrix metalloproteinase-3 is a potent prognostic factor associated with cell proliferation and migration in prostate cancer [J]. Exp Mol Pathol, 2025, 141: 104954.
[12].BALAJI S, KIM U, MUTHUKKARUPPAN V, et al. Emerging role of tumor microenvironment derived exosomes in therapeutic resistance and metastasis through epithelial-to-mesenchymal transition [J]. Life Sci, 2021, 280: 119750.
[13].FONTANA F, ANSELMI M, LIMONTA P. Adipocytes reprogram prostate cancer stem cell machinery [J]. J Cell Commun Signal, 2023, 17(3): 915-24.
[14].LIU D, WANG L, GUO Y. Advances in and prospects of immunotherapy for prostate cancer [J]. Cancer Lett, 2024, 601: 217155.
[15].FERIZ A M, KHOSROJERDI A, LOTFOLLAHI M, et al. Single-cell RNA sequencing uncovers heterogeneous transcriptional signatures in tumor-infiltrated dendritic cells in prostate cancer [J]. Heliyon, 2023, 9(5): e15694.
[16].KAUR H B, GUEDES L B, LU J, et al. Association of tumor-infiltrating T-cell density with molecular subtype, racial ancestry and clinical outcomes in prostate cancer [J]. Mod Pathol, 2018, 31(10): 1539-52.
[17].NONOMURA N, TAKAYAMA H, NAKAYAMA M, et al. Infiltration of tumour-associated macrophages in prostate biopsy specimens is predictive of disease progression after hormonal therapy for prostate cancer [J]. BJU Int, 2011, 107(12): 1918-22.
[18].WOO J R L M A, MULDONG M T, et al. Tumor infiltrating B-cells are increased in prostate cancer tissue [J]. Journal of translational medicine, 2014, 12(1): 30.
[19].XU Z C S, ZHANG Y. Roles of m5C RNA modification patterns in biochemical recurrence and tumor microenvironment characterization of prostate adenocarcinoma [J]. Frontiers in Immunology, 2022, 13: 869759.
[20].WANG M, XIA H, YAN Q, et al. Identification of Pyroptosis-Related Gene Signatures and Construction of the Risk Model to Predict BCR in Prostate Cancer [J]. Front Mol Biosci, 2022, 9: 850758.
[21].HAN S, SHI T, LIAO Y, et al. Tumor immune contexture predicts recurrence after prostatectomy and efficacy of androgen deprivation and immunotherapy in prostate cancer [J]. J Transl Med, 2023, 21(1): 194.
[22].LIU H, ZHANG W, ZHANG Y, et al. Mime: A flexible machine-learning framework to construct and visualize models for clinical characteristics prediction and feature selection [J]. Comput Struct Biotechnol J, 2024, 23: 2798-810.
[23].YOSHIHARA K, SHAHMORADGOLI M, MARTINEZ E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data [J]. Nat Commun, 2013, 4: 2612.
[24].OSORIO D, ZHONG Y, LI G, et al. scTenifoldKnk: An efficient virtual knockout tool for gene function predictions via single-cell gene regulatory network perturbation [J]. Patterns (N Y), 2022, 3(3): 100434.
[25].ZHENG R, TANG Y, DASKIVICH T J, et al. Quantifying the Impact of Biochemical Recurrence After Radical Prostatectomy on Overall Mortality by Years of Life Lost: Results from the SEARCH Database [J]. Eur Urol Oncol, 2025.
[26].PISANSKY T M, THOMPSON I M, VALICENTI R K, et al. Adjuvant and Salvage Radiotherapy after Prostatectomy: ASTRO/AUA Guideline Amendment 2018-2019 [J]. J Urol, 2019, 202(3): 533-8.
[27].LEE E E, SINGH T, HU C, et al. The impact of salvage radiotherapy initiation at PSA </= 0.5 ng/ml on metastasis-free survival in patients with relapsed prostate cancer following prostatectomy [J]. Prostate, 2023, 83(2): 190-7.
[28].KANG J H H Y S, KIM. Concern for overtreatment using the AUA/ASTRO guideline on adjuvant radiotherapy after radical prostatectomy [J]. BMC urology, 2014, 14(1): 30.

[29].KAWAMURA N, NAKAYAMA M, HAYASHI T, et al. Risk assessment of late biochemical recurrence after radical prostatectomy: Usefulness of ultra-sensitive prostate-specific antigen measurement [J]. Int J Urol, 2025, 32(4): 380-6.
[30].GATTELLI A, PELLEGRINO B, DAVID K, et al. CD74 promotes the formation of an immunosuppressive tumor microenvironment in triple-negative breast cancer in mice by inducing the expansion of tolerogenic dendritic cells and regulatory B cells [J]. PLOS Biology, 2024, 22(11).
[31].GAO P, ZOU K, XIAO L, et al. High expression of PTGES3 is an independent predictive poor prognostic biomarker and correlates with immune infiltrates in lung adenocarcinoma [J]. Int Immunopharmacol, 2022, 110: 108954.
[32].ERLANDSSON A, CARLSSON J, LUNDHOLM M, et al. M2 macrophages and regulatory T cells in lethal prostate cancer [J]. Prostate, 2019, 79(4): 363-9.
[33].FENG D, SHI X, LI D, et al. M2 macrophage-related molecular subtypes and prognostic index for prostate cancer patients through integrating single-cell and bulk RNA sequencing analysis [J]. Genes Dis, 2024, 11(4): 101086.
[34].FERNANDES S, OLIVER-DE LA CRUZ J, MORAZZO S, et al. TGF-beta induces matrisome pathological alterations and EMT in patient-derived prostate cancer tumoroids [J]. Matrix Biol, 2024, 125: 12-30.
[35].PRUITT H C, GUAN Y, LIU H, et al. Collagen VI deposition mediates stromal T cell trapping through inhibition of T cell motility in the prostate tumor microenvironment [J]. Matrix Biol, 2023, 121: 90-104.
[36].HE Z, ZHOU X, XIAO Y, et al. In vitro screening methods of novel immune checkpoint inhibitors related to T cell infiltration and anti-PD-1 resistance [J]. Methods Cell Biol, 2024, 190: 11-24.
[37].TSAI A K, KAGALWALLA S, LANGER J, et al. Pembrolizumab for metastatic castration-resistant prostate cancer: trials and tribulations [J]. Expert Opin Biol Ther, 2024, 24(1-2): 51-62.
[38].SHI Y C, AN Q, TANG N, et al. Inhibition of FAK promotes pancreatic cancer immunotherapy by mediating CXCL10 secretion to enhance CD8(+) T cell infiltration [J]. Oncoimmunology, 2025, 14(1): 2539442.
[39].CHENG B, DING K, CHEN P, et al. Anti-PD-L1/TGF-betaR fusion protein (SHR-1701) overcomes disrupted lymphocyte recovery-induced resistance to PD-1/PD-L1 inhibitors in lung cancer [J]. Cancer Commun (Lond), 2022, 42(1): 17-36.
[40].LI Y, ZHANG C, SAMAD A, et al. Structural mechanism of dsDNA recognition by the hMNDA HIN domain: New insights into the DNA-binding model of a PYHIN protein [J]. Int J Biol Macromol, 2023, 245: 125461.
[41].GU L, CASSERLY D, BRADY G, et al. Myeloid cell nuclear differentiation antigen controls the pathogen-stimulated type I interferon cascade in human monocytes by transcriptional regulation of IRF7 [J]. Nat Commun, 2022, 13(1): 14.
[42].SALMINEN A, KAARNIRANTA K, KAUPPINEN A. Myeloid cell nuclear differentiation antigen (MNDA) is an interferon-inducible regulator of myelopoiesis. Potential role in the myeloid-biased aging process [J]. Ageing Res Rev, 2026, 118: 103112.
[43].ZHANG L F, ZHANG Y, SHUI R H, et al. MNDA expression and its value in differential diagnosis of B-cell non-Hodgkin lymphomas: a comprehensive analysis of a large series of 1293 cases [J]. Diagn Pathol, 2024, 19(1): 60.
[44].JENSEN M M, BONNA A, FREDERIKSEN S J, et al. Tyrosine-sulfated dermatopontin shares multiple binding sites and recognition determinants on triple-helical collagens with proteins implicated in cell adhesion and collagen folding, fibrillogenesis, cross-linking, and degradation [J]. Biochim Biophys Acta Proteins Proteom, 2022, 1870(5): 140771.
[45].YE D, WANG Y, DENG X, et al. DNMT3a-dermatopontin axis suppresses breast cancer malignancy via

inactivating YAP [J]. Cell Death Dis, 2023, 14(2): 106.
[46].CATALáN V, DOMENCH P, GóMEZ-AMBROSI J, et al. Dermatopontin Influences the Development of Obesity-Associated Colon Cancer by Changes in the Expression of Extracellular Matrix Proteins [J]. International Journal of Molecular Sciences, 2022, 23(16).
[47].XU P, LI S, LIU K, et al. Downregulation of dermatopontin in cholangiocarcinoma cells suppresses CCL19 secretion of macrophages and immune infiltration [J]. J Cancer Res Clin Oncol, 2024, 150(2): 66.