Effect of tumor-associated macrophage on survival prognosis of prostate cancer

Yitian Ou  
Kunming Medical University Second Affiliated Hospital

Chengxing Xia  
Kunming Medical University Second Affiliated Hospital

Mingming Liu  
Kunming Medical University Second Affiliated Hospital

Haiyang Jiang  
Kunming Medical University Second Affiliated Hospital

Yong Zhu  
Kunming Medical University Second Affiliated Hospital

Delin Yang (✉ ydelin@163.com)  
Kunming Medical University Second Affiliated Hospital

Article

**Keywords:** Prostate cancer, Tumor microenvironment, Tumor-associated macrophages, Macrophage polarity, Bioinformatics, Nomogram, Immune checkpoints

**Posted Date:** July 6th, 2023

**DOI:** https://doi.org/10.21203/rs.3.rs-3107081/v1

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

Background. Prostate cancer (PCa) has become the highest incidence of malignant tumor among men in the world. Tumor microenvironment (TME) is necessary for tumor growth. This research aimed at the role of M2 macrophages polarization in PCa.

Methods. Datasets were selected from GEO and TCGA. Cluster analysis, pathway analysis were carried out. Differential-expressed-genes (DEGs) were screened and CIBERSORTx was used to calculate immune infiltration. WGCNA was used to screen the gene modules and related-genes. Cox regression analysis was carried out and redundant factors were removed. We constructed the nomogram by selecting age, T stage and M2 risk score. We further screened the DEGs between M2 high/low risk groups and analyzed the pathways, immune infiltration and most differentially expressed checkpoints between the groups.

Results. M2 macrophages promote the occurrence, development and metastases of PCa. 9 prognostic M2 related genes were screened. In M2 risk scoring model, the prognosis of high risk group was worse than low risk group. M2 risk score was an independent predictor of PCa prognosis. We further screened 34 checkpoints related to M2 risk scores.

Conclusions. M2 macrophages promote progression of prostate cancer. Based on M2 risk groups, our nomogram was proved to be effective for predicting PCa prognosis. 34 related immune checkpoints screened based on M2 risk score can be used as potential targets for follow-up study of immunotherapy.

Introduction

Prostate cancer is a malignant tumor that occurs on the epithelium of the prostate of which the most common pathological types is prostate adenocarcinoma (PRAD). According to the GLOBOCAN statistics of the World Health Organization in 2021, the incidence of PCa ranks first among male malignant tumors in the world, and its fatality rate is second only to lung cancer[1]. The clinical course of PCa patients has great heterogeneity, localized PCa patients can survive a long time without tumor progression, while invasive and metastatic PCa progress rapidly and lack of effective treatment. Androgen deprivation therapy (ADT) is an important treatment of PCa. However, almost all patients will progress to castration-resistant prostate cancer (CRPC) after 18-24 months of endocrine therapy and life quality decreased significantly. Therefore, more effective treatment and monitoring methods are urgently needed.

Including tumor and non-tumor components, TME is necessary for tumor growth. Non-tumor components include infected immune cells, blood vessels, fibrous cells and other component. The infected immune cells play an important role in the occurrence and progression of tumors. Many scholars have shown that cytotoxic T lymphocytes, mast cells and NK cells greatly effect the progression, metastasis and drug resistance of PCa[2-5].

Macrophages are important myeloid cells involved in tumor immunity. Macrophages with different polarity play different or even opposite roles. The two main polarity secretion spectra are M1 and M2.
M1 is known as anti-angiogenesis, promote chronic inflammation, inhibit tumor growth. M2 promotes blood vessel growth, organ fibrosis, tumor growth etc. In TME, M1 and M2 macrophages play opposite roles. M1/M2 ratio is an important index to evaluate the prognosis of tumors. The macrophages recruited in the TME are called tumor-associated macrophages (TAMs), which mostly show the characteristics of M2 macrophages, which makes it easier for tumor cells to escape and metastasize.

Single-cell sequencing is a new technique for sequencing transcriptome at single-cell level. The analysis based on this technique can study the gene expression in a single cell and solve the problem of cellular heterogeneity that can not be solved by previous techniques such as transcriptional sequencing, and makes it possible to analyze different components of TME and their interaction.

Immune checkpoint refers to the internal regulation mechanism of the immune system, which can maintain its own tolerance and help to avoid collateral damage during the physiological immune response. While the existing immune checkpoints are used in immunotherapy, new checkpoints are constantly being proposed. The purpose is to better monitor and put forward a new treatment therapies.

In this study, we described the immune landscape of PCa from different dimensions. By screening the M2-related genes and scoring each sample, we found that M2 macrophages promote the occurrence and progress of PCa. The M2-related genes with the highest degree of survival were screened, and the M2-related genes most related to survival were further analyzed by LASSO model. It was verified that M2 risk score was an independent predictor for PCa. Besides providing checkpoint targets for the follow-up study, our study provides a nomogram for predicting the PCa prognosis.

Materials & Methods

Data acquisition and processing

PRAD single cell transcriptome were selected from GSE137829 dataset in GEO database, including 6 samples (P1-P6). Cluster analysis was carried out by harmony and CCA. After integrating, the data were processed according to two allegation parameters: (1) To ensure the gene expression abundance and richness of each cell, the lowest threshold was 100 and the maximum threshold is unlimited. (2) The proportion of mitochondrial genes detected by each cell. The over-expression of mitochondrial genes indicates that the cells are in a "dying" state, and will lead to biased results. The threshold value was nFeature < 7500, the number of genes < 5000, and the proportion of mitochondrial genes < 15%.

The transcriptional group bulk RNA-seq is selected from the TCGA-PRAD [Log (FPKM+1)] data set as the training set. The GSE54691(104 samples) and GSE116918(248 samples) data sets are used as verification sets.

TAM screening, annotation and path analysis of scRNA-seq

The scRNA-seq data were integrated and annotated, and the cellmarker2.0 was used to label the cell subsets. The markers of M1 and M2 was extracted for follow-up analysis. The sample score of marker
gene of M1 and M2 was calculated by the PercentageFeatureSet function of R studio seurat package. Then the involved pathway analysis was done by GO and Reactome analysis.

**Screening of DEGs and calculation of immune infiltration**

The DEGs between cancer and benign tissues were screened by limma package (V3.54.1) in R studio (V4.2.1). CIBERSORTx (https://cibersortx.stanford.edu/) was used to calculate the immune infiltration.

**WGCNA analysis and screening of M2 key genes**

Weighted gene co-expression network analysis (WGCNA) was used to screen the gene modules and key genes related to M2 macrophages. The expression matrix of cancer tissue samples in TCGA-PRAD was extracted. According to the median score of M2 macrophage immune infiltration, the cancer samples were divided into high and low score group. The limma package of R Studio was used to differentiate genes between the high/low score group of M2 macrophage immune infiltration. The screening conditions were $P < 0.05$ and $\log{FC} > 0$. This method was used to find the positive correlation module which was significantly related to M2. In WGCNA, the first 25% genes were retained, and the soft threshold was screened by power threshold, according to the standard of hybrid dynamic tree cutting. Each gene module contains at least 100 genes. We drew the heat map of the correlation between the module and the shape, and selected the most significant module through the correlation coefficient and significant $P$ value of the module. Module membership (MM) and Gene significance (GS) were screened. The genes in these modules were selected for follow-up analysis. The genes selected by single cell M2 macrophage subsets markers and WGCNA were intersected by VENN map and used as M2 macrophage-related genes for follow-up research.

**Calculation of the M2 risk score and the effectiveness analysis**

Based on the screened M2 macrophage-related hub genes, the markers obtained by VENN map and scRNA-seq analysis were intersected. The cox regression model was analyzed by using the coxph function of R studio. $P < 0.1$ was used to obtain the prognostic chemokine-related genes. Then use the glmnet package of R studio to reduce overfitting, and calculate the signature score of the cases according to the gene expression level and its regression coefficient.

The formula is as follows:

$$ Score = \sum_{i=0}^{n} \beta i \times xi $$

$\beta i$ : Weight coefficient of each gene  $xi$ : Expression of each gene
Based on the signature score, cox regression analysis was carried out, and the over-fitted redundant factors were removed by LASSO regression analysis. Genes related to prognosis were obtained.

The M2 risk score of each sample is calculated by the risk model, and the median score was divided into M2 high/low risk group. The M2 risk score model was verified by training set TCGA-PRAD and verification set GSE54691,GSE116918. The KM curve, ROC curve, risk score, sample survival and the expression of prognostic factors were drawn respectively.

Age, pathological T stage and other information was collected. By combined them with M2 risk score model and used the coxph function(survival 3.5-3) for univariate and multivariate analysis to determine whether M2 risk score is an independent predictor.

**Construction of M2 prognosis-related genes nomogram**

We constructed a nomogram by selecting age, pathological T stage and M2 risk score, and took the overall survival(OS) as the observation index. According to the contribution of each influencing factor to the survival risk in the model, the accuracy was tested by calibration curve.

We used the decision curve analysis(DCA) to explore the maximum net income threshold of different variables to the survival outcome, so as to test the validity and practicability of the line chart model, and then draw the ROC curves of 2, 3 and 4 years to verify its reliability.

**HALLMARK pathway scoring and enrichment of M2 risk groups**

Taking the differential gene expression of M2 high/low group as input data, the HALLMARK pathway score of each sample was calculated by R package GSVA1.46.0, and the difference of age, T and other clinical information between M2 high and low group was shown.

**ESTIMATE scoring**

Taking the differential gene expression of M2 high/low score as input data, using ESTIMATE package to score immune infiltration. ESTIMATE was based on ssGSEA algorithm to score each sample of Stromal and Immune gene set in tumor expression matrix. Finally, the scores are combined to generate an ESTIMATE Score and calculate the tumor purity.

**Results**

**scRNA-seq reading, quality control and hypervariable gene recognition**

After preprocessing the scRNA-seq of GSE137829, we found that the data of P2 and other samples could not be clustered, so it was excluded. After combining P1, P3, P4, P5 and P6, a total of 25313 cells and 30074 genes(Supplementary Fig. 1A.B.C) were obtained. After quality control, we gained 23866 cells and
30074 genes (Supplementary Fig. 1D.E.F). After that, the relationship between gene proportion and sequence number of single cell mitochondria, and the relationship between sequence number and gene number were drawn (Supplementary Fig. 1G.H).

**Single cell principal component analysis, UMAP dimension and cell population annotation**

After normalization, principal component analysis was carried out. Samples and genes (Supplementary Fig. 2A.B) representing the characteristics of the data set were selected from the transcriptome high-throughput sequencing by dimension, and the 1st to 50th principal components were selected for follow-up analysis.

Based on the first 50 main components, the adjacent cells were determined by FindNeighbors function. The cells were grouped by FindClusters, and 34 cell groups were obtained. Then cluster analysis (Supplementary Fig. 2C.D) was carried out by UMAP.

Bubble diagram shows the expression of cell marker genes in different cell groups. Using the above methods to annotate the cell population. The markers used are basically the same as the original data set (Supplementary Fig. 3A). We got a result of 7 cell types: Basal/intermediate, Endothelial, Fibroblast, Luminal, Mast Cell, Monolytic, T Cell (Supplementary Fig. 3B). A small cell group were classified as unknown because the expression characteristics were not obvious. We selected the most significantly expressed markers, and draw its expression to determine the different cell groups (Supplementary Fig. 3C.D).

**Monolytic subgroup analysis**

After Extracting the monolytic cells, we selected the three main myeloid immune cells (monocytes, dendritic cells and TAM cells) for annotation (Supplementary Fig. 4A.B). Then we draw the marker expression highlight map and violin map (Supplementary Fig. 4C.D).

**TAM subgroup analysis**

We further explored the differences between M1 and M2 in TAM subsets. Using markers in cellmarker2.0 database, TAM was divided into M1 and M2 macrophage subsets (Supplementary Fig. 5). Cellular markers was extracted for follow-up analysis (2431 genes).

**Go and Reactome pathway enrichment**

The marker genes of M1 and M2 macrophages were analyzed by GO and Reactome pathway enrichment. In GO Biological Process (GOBP) analysis, M1 marker genes were enriched to myeloid leukocyte activation, regulation of immune effector process and other pathways (Supplementary Fig. 6A). M2 marker genes were enriched to positive regulation of cytokine production, leukocyte mediated immunity pathway (Supplementary Fig. 6E). In GO Cellular Component (GOCC) analysis, M1 and M2 marker genes were enriched to secretory granule membrane, tertiary granule and other pathways (Supplementary
Fig. 6B.F). In GO Molecular Function (GOMF) analysis, M1 and M2 marker genes were enriched to immune receptor activity, MHC protein complex binding and other pathways (Supplementary Fig. 6C.G). In the Reactome pathway enrichment analysis, M1 marker genes were enriched to Neutrophil degranulation, Interleukin-10 signaling pathway (Supplementary Fig. 6D). M2 marker genes were enriched to Neutrophil degranulation, Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell pathway (Supplementary Fig. 6H).

### TCGA DEGs screening

496 tumor samples and 52 paracancerous samples of TCGA-PRAD were analyzed to screen DEGs. The condition was that $|\log FC|>1$ and $q<0.05$. 524 DEGS were screened, of which 159 were up-regulated and 365 were down-regulated (Supplementary Fig. 7).

### CIBERSORTx of immune cells infiltration in PCa TME

The DEGs were used to calculate the immune infiltration (Supplementary Fig. 8A) using CIBERSORTx. The difference in immune infiltration between cancer and benign tissues (Supplementary Fig. 8B) was shown by box map. The results showed a significant difference in immune infiltration between cancer and benign tissues among all immune cells except T.cells.gamma.delta.

### Screening of M2 related modules and key genes by WGCNA

The DEGs were screened between high and low M2 immune infiltration groups, and a total of 4519 up-regulated genes were obtained in high groups. After WGCNA, three modules with the highest correlation with M2 macrophages (green, brown, turquoise) were selected, and 650 hub genes from these three modules were extracted for follow-up research (Supplementary Fig. 9).

### Construction and efficacy evaluation of prognostic model based on M2 marker genes

2431 M2 macrophage markers obtained by scRNA-seq were intersected with 650 M2 hub genes. A total of 12 M2 related genes were obtained by COX regression analysis. After removing redundant factors by LASSO, 9 prognostic related genes were selected: SMO2, PLPP1, HES1, STMN1, GPR160, ABCG1, MAZ, MYC and EPCAM (Supplementary Fig. 10).

### Effectiveness Analysis of M2 risk scoring model of PRAD

In the training set TCGA-PRAD, the prognosis of M2 high risk group was worse than low risk group (Fig. 1A, $P = 0.0462$). The ROC curves of 2, 3, 4 years were drawn, and the corresponding AUC values were 0.648, 0.806, 0.860 respectively (Fig. 1B). The risk score, survival time distribution and the expression of prognostic factors in the high/low score group were shown by Fig. 1, indicating that the model has a good prognostic value for the training set.

In the verification set GSE54691, the prognosis of M2 high risk group was worse than that of low risk group (Fig. 2A, $P = 0.0063$). The ROC curves 2, 3, 4 years were drawn, and the corresponding AUC values
were 0.653, 0.680, 0.684 respectively (Fig. 2B). The risk score, survival time distribution and the expression of prognostic factors in the high/low score group are shown in Fig. 2, indicating that the model has a certain prognostic value for this verification set.

In the verification set GSE116918, the prognosis of M2 high risk group was worse than that of low risk group (Fig. 3A, P = 0.02). The ROC curves of 2, 3, 4 years were drawn, and the corresponding AUC values were 0.942, 0.809, 0.658 respectively (Fig. 3B). The risk score, survival time distribution and the expression of prognostic factors in the high/low score group are shown in Fig. 3, indicating that the model has a good prognostic value for this verification set.

The univariate and multivariate analysis of age, pathological stage and M2 risk score, the results were as shown by Supplementary Table.1. Age is not an independent predictor. T staging only showed significant difference in univariate analysis (P = 0.0096). M2 risk score was an independent predictor of PCa prognosis in both univariate (P = 9.10E-08) and multivariate analysis (P = 0.0098).

Age, T stage and M2 risk score were used to construct a nomogram (Fig. 4A) for predicting OS PRAD patients. The calibration curve showed that the predicted OS was in good consistency with the actual survival time, indicating that the nomogram had a high accuracy (Fig. 4B). The reliability was verified by ROC, and the AUC values of 2, 3, 4 years were 0.703, 0.810, 0.867 respectively (Fig. 4C). DCA results showed that among all the risk factors, the model had the highest prognostic net income value, indicating that the model had good validity and practicability in predicting OS (Fig. 4D).

**DEGs screening and pathway analysis of M2 high and low risk groups**

After screening the DEGs between M2 high/low risk group, 8629 DEGs were obtained, including 4826 up-regulated genes and 3803 down-regulated genes (Supplementary Fig. 11A). The HALLMARK pathway scores of 8629 DEGs in each sample were calculated, and the pathway heat map (Supplementary Fig. 11B) was drawn. The results showed that there was a significant statistical difference between M2 high/low risk group in 41 pathways.

**Immune infiltration analysis between M2 high and low risk groups**

The results showed significant differences in ESTIMATE Score, Immune Score, Stromal Score, Tumor Purity between M2 high/low risk groups (P < 0.01, Fig. 5A.B). CIBERSORT was used to calculate the infiltration degree between M2 high/low risk groups. The results showed that there were significant differences in naive B cells, M2 macrophages and CD4 + memory T cells (P < 0.05. Figure 5C.D).

**Differentially expressed immune checkpoints between M2 high and low risk groups**
We further screened 73 immune checkpoints and found significant differences of 34 immune checkpoints between high/low M2 groups (Fig. 6).

Discussion

PCa has become a major threat to men in the world. The progression and prognosis of PCa vary greatly due to the spatial and clonal heterogeneity[6]. The development of single-cell technology provides the possibility for more accurate study of PCa transcriptome information. We depicted the immune landscape of PCa. Focusing on the analysis of the function of M2 macrophages in PCa, we screened out M2-related DEGs and the mainly involved pathways such as wnt pathway.

TME contains a variety of tumor and non-tumor components. These components interact with each other and form a dynamic network [7]. The interaction of various components in PCa TME has also become a hotspot in urology in recent years. J T W Kwon et al.[8] summarized the interaction of cytotoxic T cells, B cells, M2 macrophages and other immune cells through different cytokines in PCa. Hisham F Bahmad et al.[9] concluded that crosstalk between epithelial cells and cellular stroma plays an indispensable role in progression and metastasis in PCa. Through single cell sequencing and analysis, Chen S et al.[10] found that monocytes, dendritic cells affect tumor progression, and TME promoted the metastasis and spread of PCa when it was formed early. In this study, we found that M2 macrophage infiltration was related to the prognosis of PCa, and higher M2 score corresponds to the poor prognosis.

In clinic, the prognosis of patients with PCa is affected by great individual differences. Some patients can survive with tumors for a long time without obvious progress, while others will enter the stage of CRPC even after proper treatment. The existing clinical prognosis detection methods still have many deficiency. With the development of genomics and transcriptome research, more subtypes of PCa have been revealed, which provides the possibility for new immune checkpoints affecting the progression and metastasis of PCa[11–12]. Y He et al.[13] concluded that MMR gene deficiency can enhance PRAD anti-tumor immune response, showing more tumor infiltrating lymphocytes. Some mCRPC patients will have CDK12 aberration, and the inactivation of CDK12 will lead to focal tandem repetition, which will enhance tumor immune response. Abida et al.[14] treated 11 mCRPC patients with metastatic high micro-satellite instability or mismatch repair defect with immune checkpoint inhibitors, of which 6(54.5%) had PSA decreased by more than 50%. In our study, nine prognostic related genes were screened, and a nomogram prognostic prediction model was constructed, which showed a good ability to predict prognosis in both training and verification set. The predicting results was consistent with clinical prognosis. We further confirmed that M2 score was an independent predictor of PCa prognosis.

Different from tradition therapies, tumor immunotherapy takes effect by activating immunity. Nowadays, there are common immune checkpoint inhibitor targets such as PD-1/PD-L1, CTLA-4. Studies have shown that CTLA-4 inhibitor Epimazumab has a certain effect on early PCa patients, but no improvement in OS in CRPC patients[15]. Yang JC et al. Found that PD-1/PD-L1 inhibitors are poor-effected when used alone in lung cancer[16]. All these suggest that immunotherapy still face great challenges. In this study, we
found 34 immune checkpoints between high/low M2 groups, which may provide targets for follow-up research.

Yet, there are still limitations in our study. Due to the high cost, single cell sequencing can’t be widely used currently. However, with the continuous improvement of single cell sequencing, this study can provide immune checkpoints for PCa patients. Besides, although the nomogram shows good prognostic value, it has not yet been verified by clinic. This will be verified in our later studies.

Conclusions

In prostate cancer, it is further verified that M2 macrophages promote the occurrence, development and metastases of PCa. M2 macrophage-related differential genes are mainly involved in tumor-related pathways such as wnt pathway. Risk score model and nomogram have good ability to predict the survival and prognosis of PCa. M2 risk score was an independent predictor of PCa prognosis.

Declarations

Data availability


References


Supplementary Figures

Supplementary Figures are not available with this version

Supplementary Table

Supplementary Table 1 is not available with this version

Figures
Figure 1

M2 risk scoring model's prognostic value of training set TCGA -PRAD. (A) KM curve of TCGA training set. The survival prognosis of high score group was significantly lower than that of low score group (P = 0.0462). (B) ROC curves of 2, 3, 4 years. (C) Risk score results. (D) Survival time distribution. (E) Expression of prognosis-related genes.
Figure 2

M2 risk scoring model's prognostic value of verification set GSE54691. (A)KM curve of GSE54691 verification set. The survival prognosis of high score group was significantly lower than that of low score group (P=0.0063). (B)ROC curves of 2,3,4 years. (C)Risk score results. (D)Survival time distribution. (E)Expression of prognosis-related genes.
Figure 3

**M2 risk scoring model’s prognostic value of verification set GSE116918.** (A) KM curve of GSE116918 verification set. The survival prognosis of high score group was significantly lower than that of low score group (P = 0.02). (B) ROC curves of 2, 3, 4 years. (C) Risk score results. (D) Survival time distribution. (E) Expression of prognosis-related genes.
Figure 4

**PRAD OS predicting nomogram of M2 risk score.** (A) Predictive OS nomogram of PRAD patients. (B) The consistency curve of actual OS and predicted OS in 2,3,4 year. (C) ROC curves of 2,3,4 years and corresponding AUC. (D) The DCA results showed the net income values of different prognostic factors, and complex represents the complex of the factors in the model.
Figure 5

**Immune cells infiltration analysis.** (A) Heat map and (B) Box diagram of ESTIMATE Score Immune Score Stromal Score Tumor Purity between M2 high/low risk groups. (C) Heat map and (D) Box diagram of immune infiltration between M2 high/low risk groups.
**Figure 6**

**Immune checkpoints analysis.** (A) Box diagram and (B) Heat map of differentially expressed immune checkpoints between M2 high/low risk groups.