Am J Cancer Res 2023;13(7):2861-2877
www.ajcr.us /ISSN:2156-6976/ajcr0150097

Original Article

Novel gene signature for predicting biochemical recurrence-free survival of prostate cancer and PRAME modulates prostate cancer progression

Gaohaer Kadeerhan[1*], Bo Xue[1,2*], Xiaolin Wu[1], Xiaofeng Hu[1,2], Jun Tian[1], Dongwen Wang[1]

[1]National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen 518116, China; [2]Shanxi Medical University, Shanxi 030012, China. [*]Equal contributors.

Received March 12, 2023; Accepted June 2, 2023; Epub July 15, 2023; Published July 30, 2023

Abstract: Biochemical recurrence (BCR) is considered as an early sign of prostate cancer (PCa) progression after initial treatment, such as radical prostatectomy and radiotherapy; hence, it is important to stratify patients at risk of BCR. In this study, we established a robust 8-gene signature (APOF, Clorf64, RPE65, SEMG1, ARHGDIG, COMP, MKI67 and PRAME) based on the PCa transcriptome profiles in the Cancer Genome Atlas (TCGA) for predicting BCR-free survival of PCa, which was further validated in the MSK-IMPACT Clinical Sequencing Cohort (MSKCC) PCa cohort. Moreover, we found that one risk-related gene (PRAME) was upregulated in tumor samples, particularly in high-risk group was well as in patients metastatic tumor and was correlated with chemotherapeutic drug response. In vitro experiments showed that knocking down PRAME reduced the proliferation, migration, and invasion of PCa cells. Therefore, our study established a new 8-gene signature that could accurately predict the BCR risk of PCa. Inhibition of PRAME attenuated the proliferation, invasion, and migration of PCa cells. These findings provide a novel tool for stratifying high-risk PCa patient and shed light on the mechanism of PCa progression.

Keywords: Prostate cancer, biochemical recurrence, prognostic signature, PRAME

Introduction

Prostate cancer (PCa) is the most common cancer in men worldwide [1]. According to the latest cancer statistics, PCa has overtook the lung cancer and becomes the highest incidence malignant tumor and second leading cause of cancer death in the United States [2]. The incidence of PCa has been significantly rising in recent years due to the increasing use of prostate-specific antigen (PSA) detection in PCa low-countries including China [3, 4].

Although radical prostatectomy or radiotherapy has demonstrated an excellent long-term control of PCa [5], approximately 20-40% PCa patients will experience biochemical recurrence (BCR) within 10 years after initial therapy [6, 7]. BCR indicates the regrowth of prostate cancer cells and is closely related to the subsequent progression to clinical metastasis that irreversibly results in patient death [7]. Thus, a better understanding of the BCR risk and exploring valuable predictive biomarkers for the prognosis of PCa are urgently needed.

BCR is defined as a detectable rising serum PSA level after local therapy [8]. Current BCR risk assessment is often based on clinicopathological characteristics such as Gleason score and pathological tumor stages, alone or in combination [8, 9]. However, due to the complex heterogeneities of PCa, patients with same clinical features may develop different outcomes, which limits the application of conventional clinical parameters for predicting patients' prognosis.

Deep sequencing technology in recent years has revealed numerous DNA-based or RNA-based molecular signatures that are involved in genetic and epigenetic alterations of PCa, which may serve as potential prognostic factors for PCa [10]. Although efforts have been made to construct gene signatures for BCR patients at high risk from defined sets of genes related to

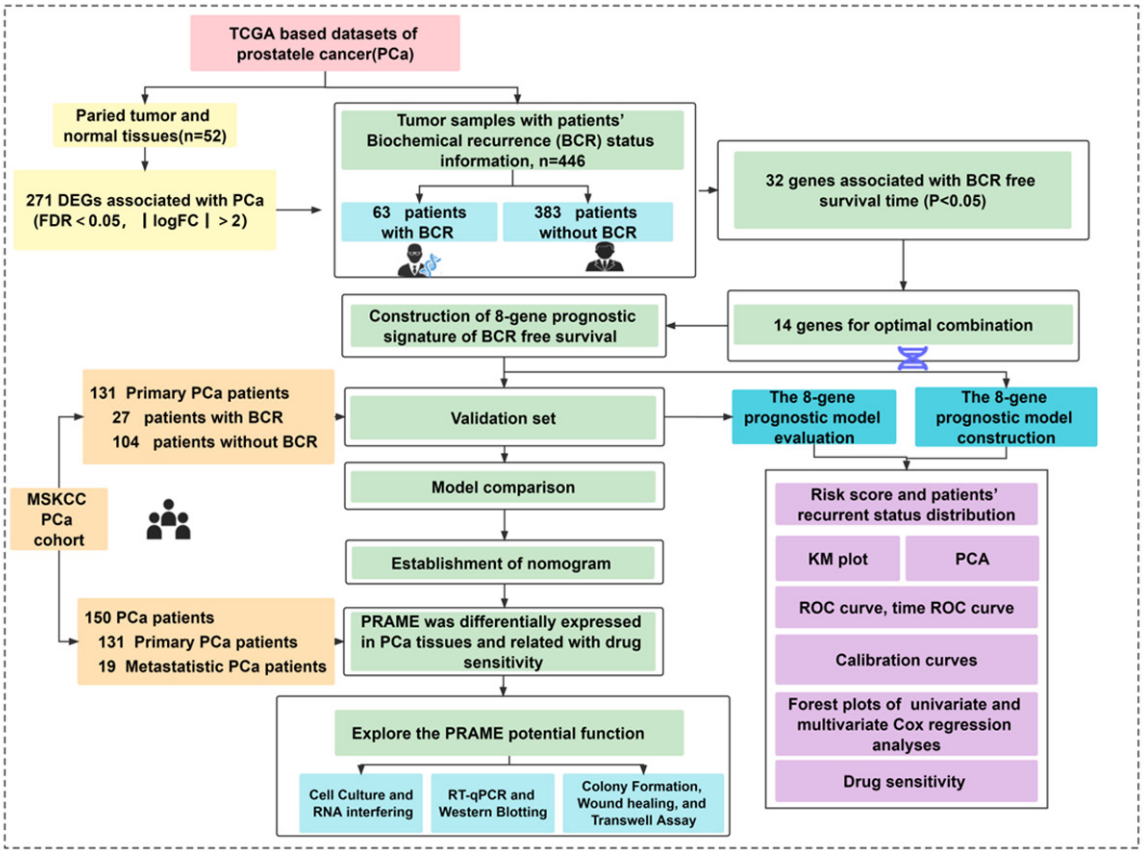


Figure 1. Flowchart of the study design. BCR, biochemical recurrence; MSKCC, MSK-IMPACT Clinical Sequencing Cohort; PCa, prostate cancer; TCGA, The Cancer Genome Atlas.

specific cellular hallmarks of cancer [10-14], it still remains challenging to accurately predict the risk of BCR due to the complicated molecular mechanisms involved in PCa relapse and progression. In addition, only a limited number of reported genes have been further functionally validated [15]. Therefore, it is imperative to identify and construct robust gene signatures related to BCR-free survival and explore the potential role of these genes in PCa progression.

In this study, we comprehensively explored the gene signatures for BCR-free survival of PCa and established a prognostic signature comprised of eight genes, as well as experimentally investigated the function of one risk-related gene in PCa progression.

Material and methods

Study cohort

The study design and workflow are presented in **Figure 1**. The RNA sequencing data of 499 PCa patients and 52 normal samples with corresponding clinical information in the TCGA-PRAD (prostate adenocarcinoma) were downloaded from the TCGA database (https://portal.gdc.cancer.gov/, accessed May 25, 2021) and used as the discovery set in this analysis. Among the 499 PCa patients, 446 have reported BCR information and BCR-free survival time. For the validation set, the gene expression profiles and the corresponding clinical data of 150 PCa patients in the MSK-IMPACT Clinical Sequencing Cohort (MSKCC) PCa cohort were downloaded from the Cbioportal database (http://cbio.mskcc.org/cancergenomics/, GSE-21032, accessed Jun 16, 2022). Of the 150 PCa patients, 131 have complete BCR information. The clinical features of these two data sets are summarized in Tables S1, S2.

Identification of genes associated with PCa

We screened mRNAs associated with PCa in the discovery set. For mRNA expression data, we removed mRNAs that have more than seventy-five percent of values as zero and kept

16938 mRNAs. Differentially expressed genes (DEGs) were selected in 52 paired PCa and paracancerous tissues using "DESeq2" package based on R software with the corresponding false discovery rate (FDR) <0.05 and log2|fold change (FC)| >2.0.

Construction and validation of PCa prognostic signature

The prognostic signature was constructed based on 446 PCa patients with BCR information in the discovery set. First, DEGs associated with BCR-free survival (P<0.05) were identified by univariate Cox regression analyses. Then, LASSO analysis was performed to select the genes predictive of BCR by using the "glmnet" package. We further narrowed down the BCR-free survival related genes by backward stepwise Cox regression analyses to optimize the prognostic model. By log-rank analysis, the association between the final selected genes and BCR-free survival was tested and visualized using Kaplan-Meier (K-M) curves via "survminer" and "survival" packages, respectively.

The prognostic signature was built with these genes, and the risk score was calculated using the formula:

Risk score = β1 × X1 + β2 × X2 + β3 × X3.....βn × Xn.

Regression coefficient (βn) of each gene derived from the multivariate Cox proportional hazard analysis in TCGA cohort was used as β score. Xn represented the standardized expression value of each gene. We detected the best cut-off for the 8-gene signature in TCGA cohort using the R "survminer" package, and the PCa patients were assigned to low- and high-risk groups by the cut-off score of 5.700.

To evaluate the prognostic signature, after plotting the risk score curve and BCR status distribution, the K-M curves and log-rank test were performed to compare the BCR-free survival between the two risk groups. Principal component analysis (PCA) was used to demonstrate the risk group distribution based on "pca3d" package. Through "pROC" and "timeROC" packages, the receiver operating characteristic (ROC) and the time-dependent ROC analysis were used to respectively test the model's discriminative ability for BCR status and the prediction accuracy for 1-, 3-, 5- and 8-year BCR-free survival. The calibration curves were performed and visualized for 1-, 3-, 5- and 8-year BCR-free survival by the "rms" package to estimate the consistency between the actual and the predicted BCR-free survival probabilities. Univariate and multivariate Cox regression analyses were used to explore whether the risk score could be an independent prognostic index in PCa. The relationship between the gene signature and clinical features was also explored. The model was further verified in the validation set using the same analyses as described above. Furthermore, the entire ROC curves and the AUCs at 1-, 3-, 5- and 8-year for BCR risk of our risk score model were compared to two published transcriptomic signature [11, 12] by "pROC" package and "survcomp" R packages.

Establishment and assessment of the nomogram

In both the discovery and validation sets, we built a nomogram for predicting 1-, 3-, 5-, 8-year BCR-free survival based on the results of the multivariate Cox regression analysis of the risk score and patient clinical features including age, Gleason score, and pathological T stage. The consistency index (C index), time-dependent ROC, and calibration curve were used to evaluate the predictive performance of the nomogram.

Drug sensitivity analysis

The response of each PCa patient in both the discovery and the validation cohorts to abiraterone, docetaxel, olaparib and bicalutamide were obtained from the Broad Institute's Cancer Therapeutics Response Portal (CTRP) [16] and Sanger's Genomics of Drug Sensitivity in Cancer (GDSC) [17] using the "oncoPredeict" R package [18]. Half-maximal inhibitory concentration (IC50) was calculated to measure the response of tumor cells to drugs.

Cell lines and RNA interfering

Human PCa cell lines (DU145 and PC-3) were obtained from the Procell Life Science & Technology (Wuhan, China) and were cultured in RPMI-1640 medium, supplemented with penicillin (100 U/ml)/streptomycin (100 μg/ml)

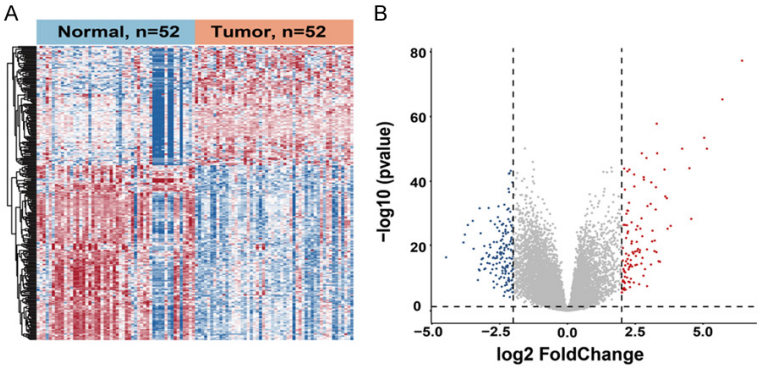


Figure 2. Initial screening of PCa-related genes in the TCGA cohort. (A) Heat map and (B) volcano plot of the DEGs between tumor and paired normal tissues. DEG, differentially expressed genes; PCa, prostate cancer; TCGA, The Cancer Genome Atlas.

and 10% fetal bovine serum in a humidified incubator with 5% CO_2 at 37°C. The lentiviral constructs expressing PRAME shRNAs (shPRAME-1: GCTGGACTCTATTGAAGATTT, shPRAME-2: GCCAGATGATTAATCTGCGTA, shPRAME-3: CCTGTGATGAATTGTICTCCT, shPRAME-4: GCTCCCAGCTTACAACCTTAA) or the non-specific control shRNA were synthesized by Tsingke Biotechnology Co., Ltd. (Beijing, China).

Real time quantitative PCR (RT-qPCR) and western blotting

Standard RT-qPCR protocol was followed. Briefly, total RNA was extracted from cells using Trizol reagent (Invitrogen, Waltham, MA, United States). Approximately 500 ng of RNA was used for the reverse transcription reaction with PrimeScript RT Master Mix [Takara Biotechnology (Dalian) Co., Ltd., China]. RT-qPCR was performed using Premix Ex TaqTM II [Takara Biotechnology (Dalian) Co., Ltd.] with the ABI7500 Real-Time PCR system (Thermofisher, USA). GAPDH was used as the internal control. The primer sequences are: PRAME forward: CAGGACTTCTGGACTGTATGGT, reverse: CTACGAGCACCTCTACTGGAA; GAPDH forward: ATCATCAGCAATGCCTCCTG, reverse: ATGGACTGTGGTCATGAGTC.

For western blot analysis, cells were lysed with RIPA buffer, and the cell lysates were separated on SDS-PAGE gel and transferred onto PVDF membranes, followed by incubation with primary antibodies at 4°C overnight. After extensive washing, the membranes were incubated with HRP-conjugated secondary antibodies, and the protein signals were detected by ECL using Bio-Rad ChemiDoc XRS+ (Bio-Rad, USA).

Colony formation, wound healing, and transwell assay

For colony formation assay, shRNA transfected DU145 and PC-3 cells in the logarithmic growth stage were seeded at 500 cells/well into 6-well plates and cultured for 7 days. Cell culture medium was changed every 3 days. The cell clones were fixed with 4% paraformaldehyde and stained with 0.5% crystal violet solution. The number of colonies was counted under a microscope.

For wound healing assay, PCa cells (1×10^5-1×10^6 cells/well) were seeded into 6-well plates. Wound scratching was created by a 100 μL pipette. The wound area was photographed at 0 h, 12 h and 24 h by time-lapse microscopy, and the migrated cells were calculated by the closure of the wound area.

For cell invasion assay using transwell, the upper surface of the membrane were coated with matrigel (Matrigel™ GFR Membrane Matrix, USA) before use. PCa cells (1×10^5 cells/mL) were seeded into the upper chamber, while the complete medium was added to the lower well. After 24 h incubation, the cells in the upper chamber were removed with a cotton swab. The cells invaded to lower surface of the membrane were fixed with 4% paraformaldehyde, stained with crystal violet solution, and counted under a microscope.

Statistical analysis

All statistical analyses were performed by R software (version 4.0.2). The gene expression comparisons across groups were performed using Wilcoxon Rank Sum and Signed Rank Tests, and the results were visualized by box plots. All statistical tests were two-sided, and $P<0.05$ was considered statistically significant.

Results

DEGs between paired PCa and normal samples

The expression profile of 52 PCa and their paracancerous tissues in TCGA cohort was com-

pared, and 271 DEGs were identified at FDR<0.05, among which, 161 was downregulated, while 110 was upregulated in tumor samples (Table S3). The heatmap and volcano plot of these DEGs were presented in **Figure 2A**, **2B**.

Construction of prognostic gene signature based on DEGs

PCa patients with information on BCR status and BCR-free survival in TCGA (n=446) were used to identify BCR-free survival related genes. The univariate Cox regression analysis revealed that 32 DEGs were significantly associated with BCR-free survival (P<0.05, **Figure 3A**). By LASSO Cox regression analysis, we excluded the colinearity of these genes and obtained 14 prognosis-related genes (**Figure 3B**). To further optimize the BCR-free survival-related genes, backward stepwise Cox regression analyses were performed, and finally 8 genes were selected (Table S4). The K-M analysis showed that the low expression of APOF, Clorf64, RPE65 and SEMG1, while the high expression of ARHGDIG, COMP, MKI67 and PRAME was significantly associated with poor BCR-free survival (**Figure 3C**, all P<0.05).

Based on the mRNA expression levels of these 8 genes, a risk score was calculated for each PCa patient using the formula: risk score = (-0.13852 × APOF) + (0.09994 × ARHGDIG) + (-0.01310 × C1orf64) + (0.21306 × COMP) + (0.18928 × MKI67) + (0.05496 × PRAME) + (-0.03966 × RPE65) + (-0.04365 × SEMG1), with a higher score indicating worse survival potential.

Evaluation and validation of PCa prognostic gene signature

The predictive performance of this 8-gene signature was evaluated in the discovery set (the TCGA cohort). The PCa patients were divided into a low-risk group (n=295) and a high-risk group (n=151) based on the best cut-off score (5.700). As shown in **Figure 4A**, **4B**, when patients were ranked according to their risk scores, the number of patients with BCR event was increased along with the increasing risk score. The K-M analysis showed that patients in the high-risk group had a higher probability of BCR compared to those in the low-risk group (P<0.0001, **Figure 4C**). In addition, the patients in the high-risk group and low-risk group presented distinction three-dimensional spatial distributions in the PCA analysis (**Figure 4D**). The ROC analysis demonstrated the risk score was able to discriminate BCR status with an AUC of 0.743 (95% confidence interval (CI): 0.681-0.806, P<0.001, **Figure 4E**). Furthermore, the time-dependent ROC analysis showed that the predictive accuracy of this gene signature for 1-, 3-, 5-, and 8-year BCR-free survival were 0.727, 0.736, 0.778 and 0.778 (**Figure 4F**), respectively. Moreover, the calibration plot indicated an outstanding consistency between the predicted and actual value (**Figure 4G**), as the C-index of the gene signature was 0.727 (95% CI: 0.659-0.796) and the bias-corrected C-index with 1000 bootstrap replications was also 0.727.

We further validated the predictive power of this 8-gene signature in the validation cohort (the MSKCC cohort, n=131), in which, 85 were assigned in the low-risk group and 46 in the high-risk group based on the risk score (1.709). Consistently, the risk score was positively associated with BCR (**Figure 5A**, **5B**) and poor BCR-free survival (P=0.024) (**Figure 5C**). These two risk subgroups also presented different distributions in the three-dimensional plane (**Figure 5D**). In addition, this 8-gene signature could accurately distinguish BCR status with AUC of 0.671 (95% CI: 0.554-0.788, P=0.003, **Figure 5E**) and predict the 1-, 3-, 5-, and 8-year BCR-free survival (0.748, 0.717, 0.719, and 0.637, respectively) (**Figure 5F**). Calibration plot displayed the consistency between the predicted and actual value in the 1-, 3-, 5-, and 8-year BCR-free survival (**Figure 5G**). The C-index of the gene signature was 0.694 (95% CI: 0.590-0.798) and the bias-corrected C-index with 1000 bootstrap replications was 0.695. Furthermore, there was no difference in the AUCs of BCR prediction between the discovery and the validation cohorts, suggesting the reliability of this 8-gene signature in predicting the BCR risk.

The 8-gene signature was an independently prognostic index in PCa

The univariate Cox analysis revealed that the risk score was significantly associated with BCR-free survival, with a hazard ratio (HR) of 3.46 (95% CI: 1.93-6.19, P<0.001) in the dis-

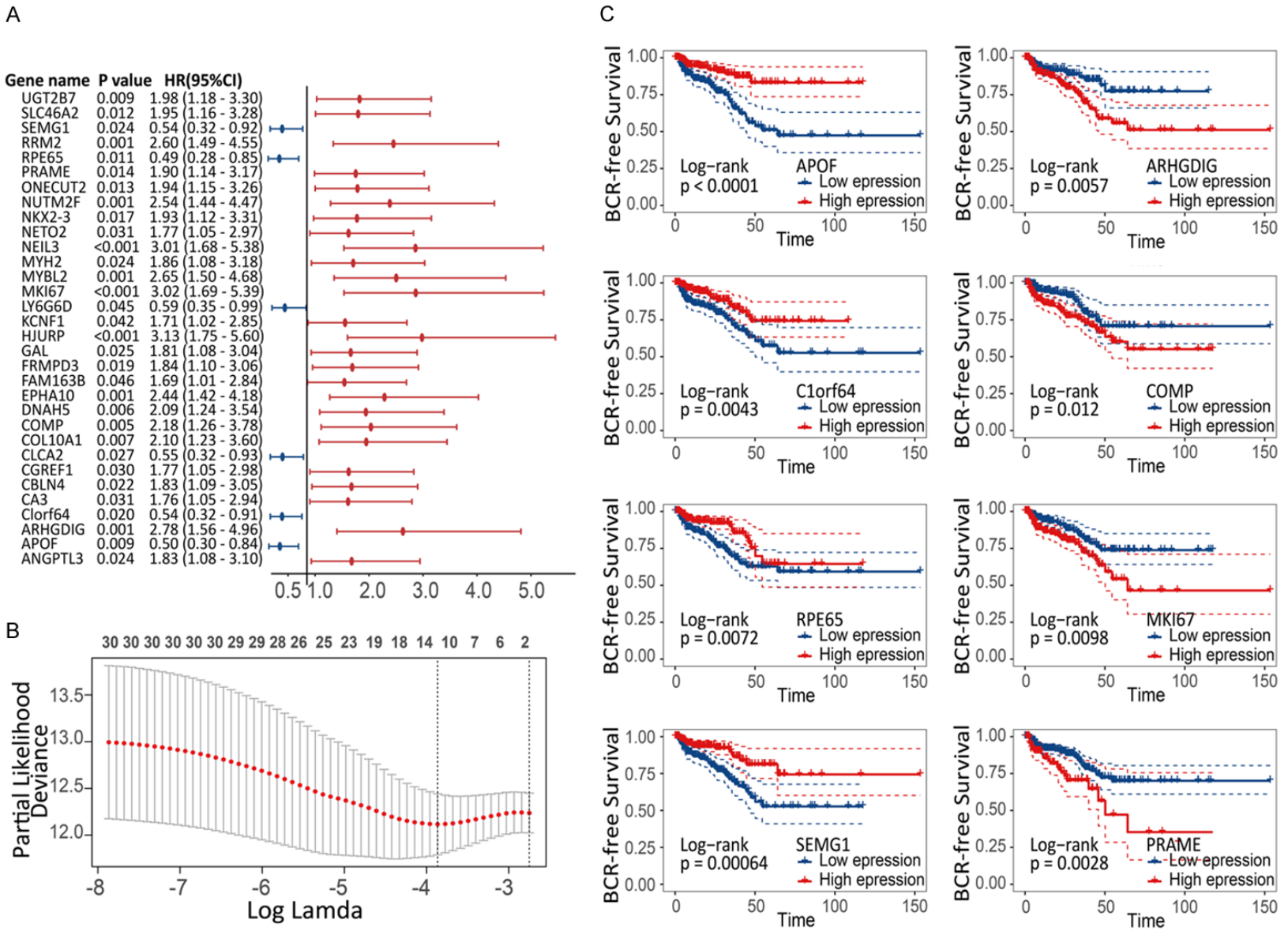


Figure 3. Identification of genes associated with the BCR-free survival of PCa patients. A. The hazard ratios of 32 BCR related genes in univariate Cox regression model. B. Partial likelihood deviance of the different number of variables revealed by the LASSO regression model. C. The Kaplan-Meier curves of eight genes (APOP, ARHGDIG, C1orf64, COMP, RPE65, MKI67, PRAME, and SEMG1) selected by stepwise Cox regression model. BCR, biochemical recurrence; PCa, prostate cancer.

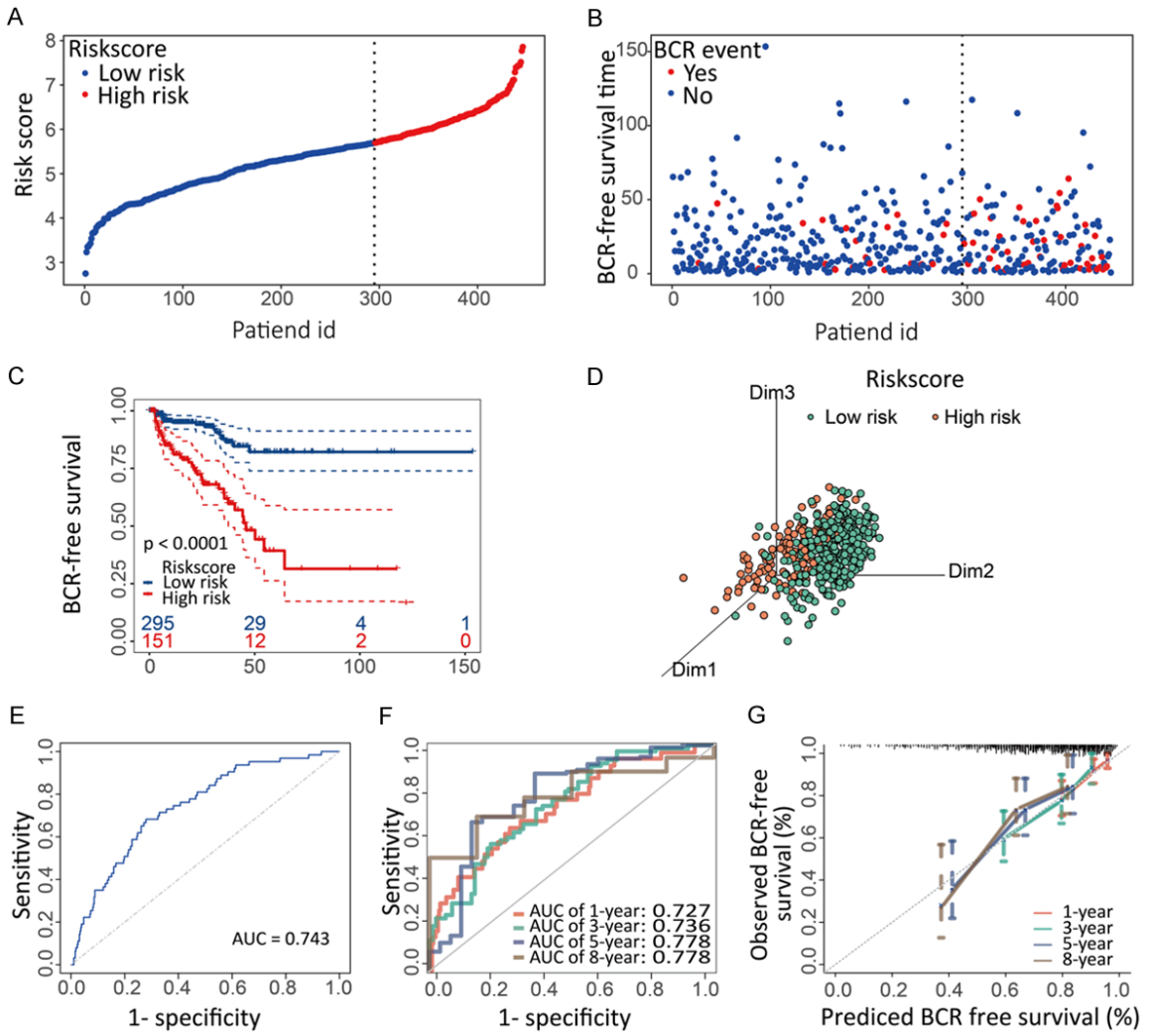


Figure 4. Prognostic model construction for BCR-free survival of PCa based on the eight-gene signature in TCGA cohort. (A) The risk score curve and (B) the distribution of BCR status. (C) K-M curves of BCR free survival of the low and high-risk groups stratified based on the cut-off scores (5.700). (D) PCA of the two risk groups. ROC analysis based on risk scores (E) for the discrimination ability of BCR status and (F) for the prediction of the 1-, 3-, 5-year and 8-year BCR free survival. (G) Calibration curves of the risk score for predicting the 1-, 3-, 5-, and 8-years BCR free survival. AUC, area under the receiver operating characteristic curve; BCR, biochemical recurrence.

covery cohort (**Figure 6A**) and of 2.84 (95% CI: 1.30-6.22, *P*=0.009, **Figure 6C**) in the validation cohort. Multivariate Cox regression analysis adjusted for age, Gleason score, and pathological T stage further demonstrated that this 8-gene signature was an independent prognostic index in both cohorts (**Figure 6B** and **6D**), with HRs of 2.29 (95% CI: 1.21-4.33, *P*=0.011) and 2.28 (95% CI: 1.01-5.12, *P*=0.046), respectively.

Model comparison

More importantly, we compared the predictive performance of this 8-gene signature with two published signatures [11, 12] and found a superior performance of our signature. In the TCGA cohort, the AUC for BCR status was 0.743 in our model, while it was 0.682 in Zhao's study and 0.598 in Hu's study (Figure S1A). The AUCs for BCR risk at 1-, 3-, 5-, and 8-year were 0.727, 0.736, 0.778 and 0.778, respectively, in our signature, while they were respectively 0.678, 0.639, 0.641 and 0.665, or respectively 0.585, 0.571, 0.540, and 0.507, in the other two studies (Figure S1C). In the MSKCC cohort, the AUCs for BCR status were respectively 0.671, 0.551, and 0.547 in ours, Zhao's and Hu's study (Figure S1B). The AUCs for BCR risk at 1-, 3-, 5-, and 8-year in our study were respectively

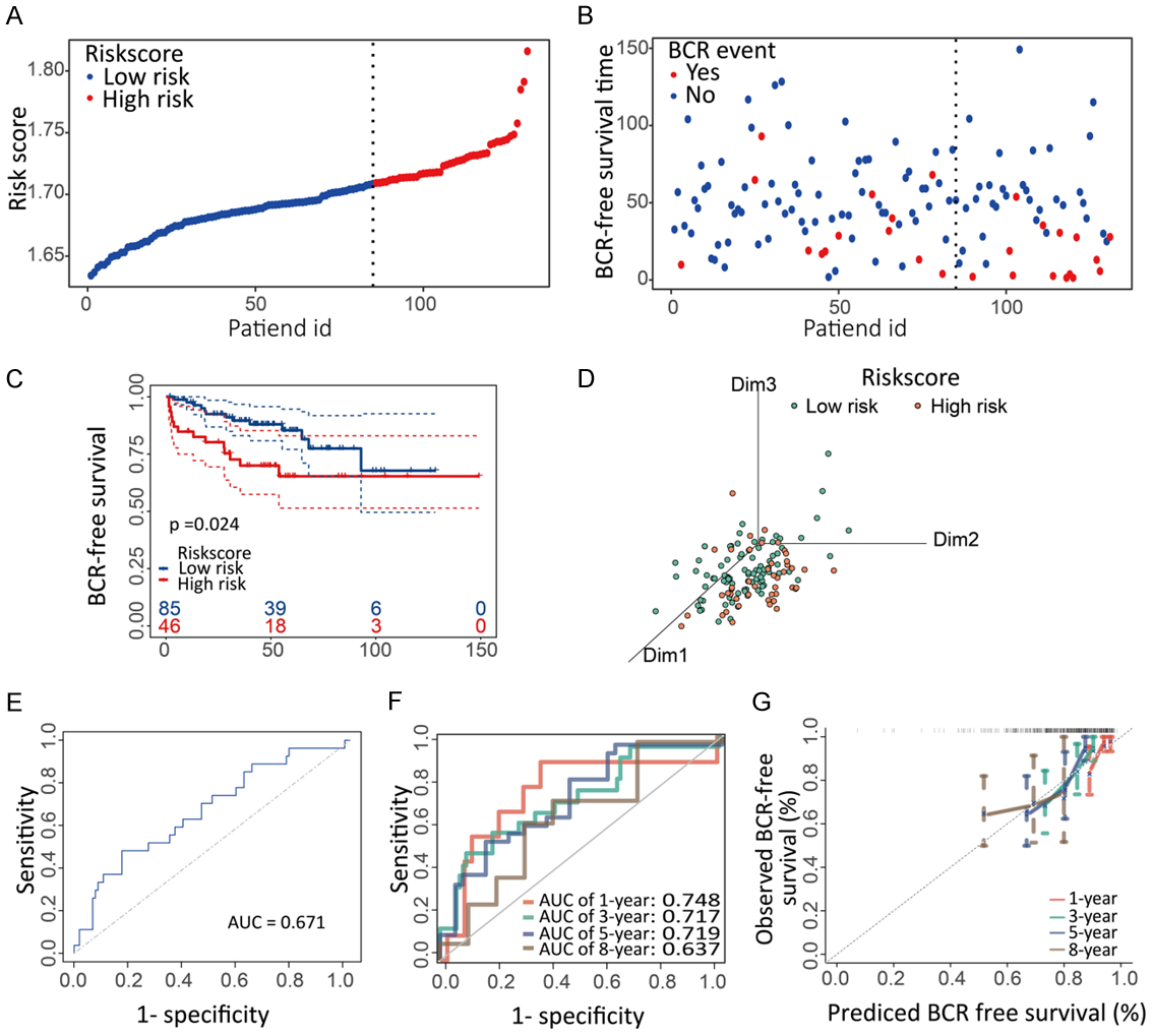


Figure 5. Validation of the prognostic signature in validation set (the MSKCC cohorth, GSE21032). (A) The curve of risk score and (B) the distribution of BCR status of the patients. (C) K-M curves of BCR free survival in the low- and high-risk groups stratified based on the calculated risk score (1.709) same as percent rank of the modeling cohort. (D) PCA of the two risk groups. ROC analysis based on risk scores (E) for the discrimination ability of BCR status and (F) for the prediction of the 1-, 3-, 5-year and 8-year BCR free survival. (G) Calibration curves of the risk score for predicting the 1-, 3-, 5-, and 8-years BCR free survival. AUC, area under the receiver operating characteristic curve; BCR, biochemical recurrence.

0.748, 0.717, 0.719 and 0.637, while they were respectively 0.693, 0.604, 0.614 and 0.586 in Zhao's study and were respectively 0.688, 0.546, 0.629, and 0.471 in Hu's study (Figure S1D). Together, these results demonstrated that our signature performed better than the other two published signatures (P<0.001).

Establishment and assessment of the nomogram

Since clinical features including T stage and Gleason score are associated with BCR-free survival, we integrated the risk score with pathological T stage as well as Gleason score to establish a nomogram in both cohorts. A nomogram score for the 1-, 3-, 5-, and 8-year BCR-free survival were obtained based on factors included in the nomogram (**Figures 7A** and S2A). The C-indexes showed the good predictive accuracy of this nomograph in these two cohorts (0.754 for the discovery cohort and 0.789 for the validation cohort). The calibration curves were also consistent in the 1-, 3-, 5- and 8-year BCR-free survival between the predicted and actual value (**Figures 7B** and S2B).

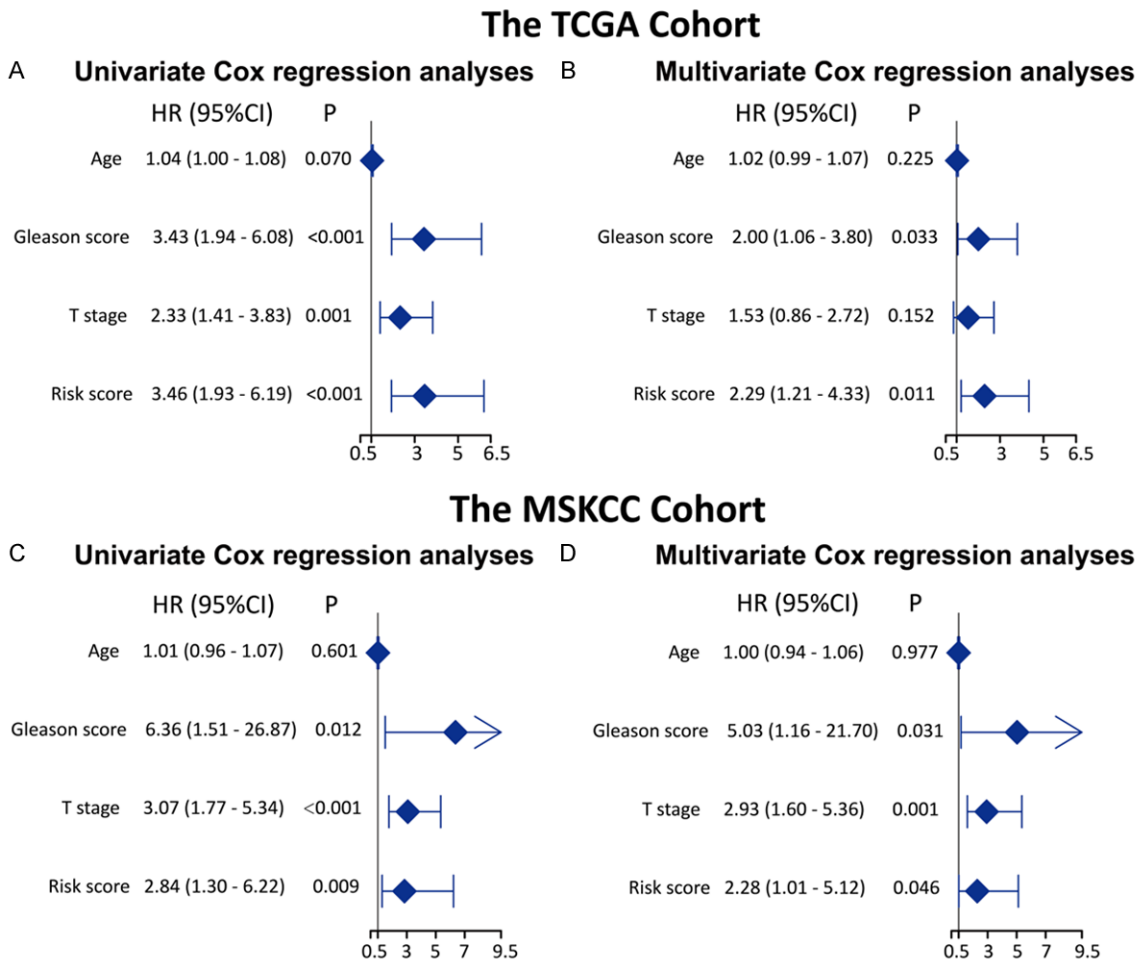


Figure 6. Cox regression analyses of clinical characteristics and risk score model. Univariate and multivariate Cox analyses of the age, Gleason score, tumor T stage, risk score in the (A, B) TCGA cohort, and the (C, D) validation MSKCC cohort. BCR, biochemical recurrence; PCa, prostate cancer.

Furthermore, the time-dependent ROC analysis indicated that the nomogram scores yielded higher AUCs in predicting 1-, 3-, 5-, and 8-year BCR-free survival than other factors, with respectively 0.746, 0.771, 0.805 and 0.808 in the discovery cohort and respectively 0.885, 0.866, 0.806 and 0.678 in the validation cohort (**Figures 7C** and S2C). The nomogram performance was similar between these two cohorts (P=0.621).

Drug sensitivity analysis of PCa patients

The predicted IC50 value showed that the high-risk group in both cohorts had a better response to olaparib (p values were <0.0001, **Figure 8A**, **8B**), indicating that olaparib may be more beneficial to patients in the high-risk group than in the low-risk group. Similarly, bicalutamide and docetaxel showed a better response in the high-risk group of discovery cohort (**Figure 8A**), though the difference was not statistically significant in the validation cohort (P=0.310 and P=0.170, respectively, **Figure 8B**). In contrast, patients in the two risk groups of both cohorts did not showed consistent response to abiraterone (**Figure 8A**, **8B**).

PRAME was differentially expressed in PCa tissues and was related with drug sensitivity

To explore the molecular mechanism involved in this risk signature, we evaluated the expression of these signature genes and found the expression of ARHGDIG, COMP, MKI67, and PRAME were upregulated, while the expression of RPE65 and SEMG1 were downregulated in the tumor samples compared to the normal samples, as well as in the high-risk group compared to the low-risk group in both cohorts

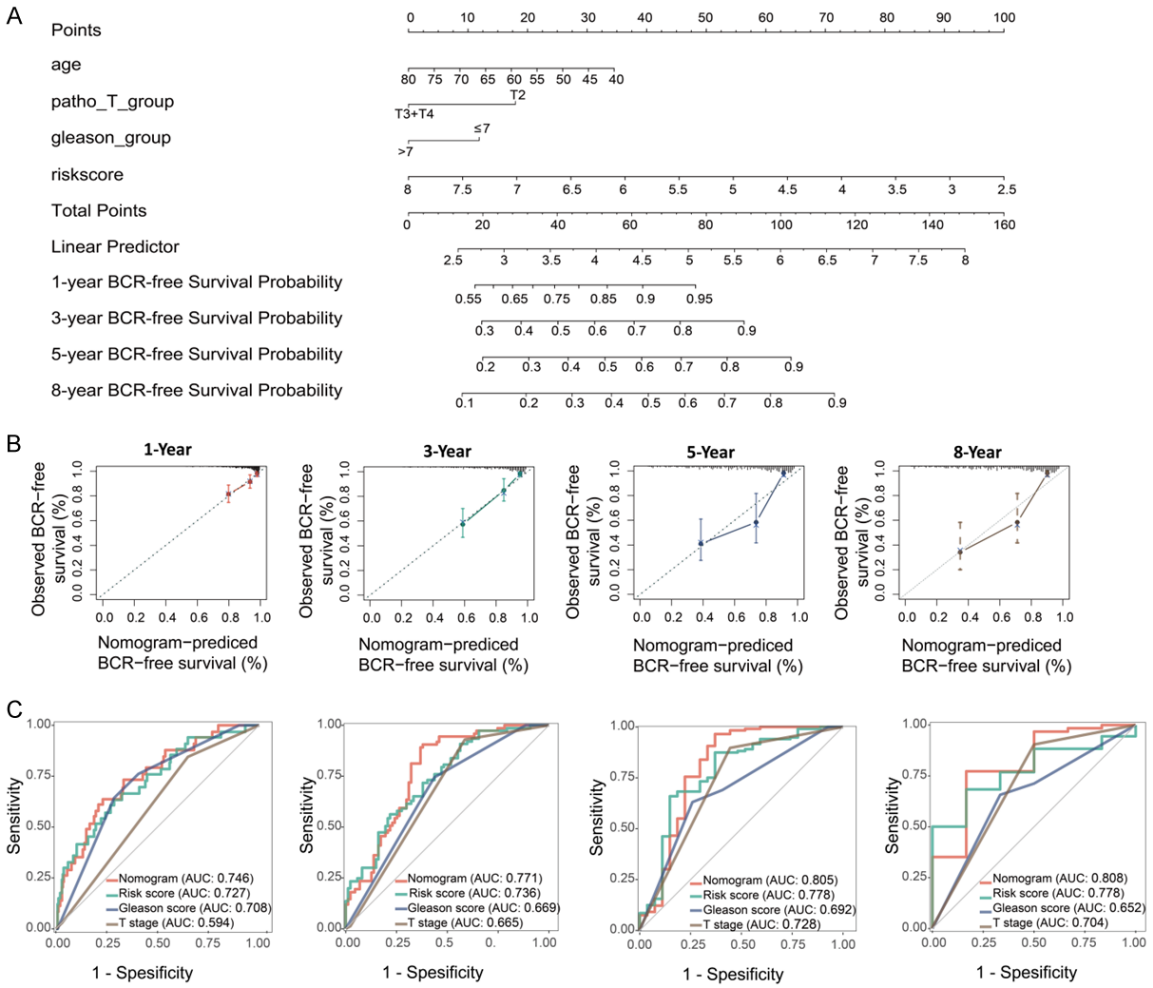


Figure 7. Nomogram for predicting the BCR-free survival of PCa patients in the TCGA cohort. A. A prognostic nomogram including signature risk score and clinical factors. B. The calibration curves of the 1-, 3-, 5-, and 8-year BCR-free survival. C. ROC curve used to evaluate the predictive performance of 1-, 3-, 5-, and 8-year BCR-free survival. BCR, biochemical recurrence; PCa, prostate cancer.

(**Figure 9A**, **9B**). Among them, the expression of PRAME was significantly correlated with olaparib, docetaxel and bicalutamide responses in both two cohorts (**Figures 9C** and S3A, S3B). More importantly, PRAME expression was elevated in the metastatic tumor sample of the validation cohort (Figure S4A) as well as in most of TCGA cancers (Figure S4B), suggesting the prognostic value of PRAME; hence, we focused our experimental investigation on PRAME.

Knockdown of PRAME inhibited PCa cell proliferation, migration, and invasion

We first verified the upregulated expression of PRAME in PCa tumor tissues by IHC in the Human Protein Atlas (**Figure 10A**). Then, we explored the effect of PRAME knockdown by shFRAME function on the proliferation, migration, and invasion of DU145 and PC3 cells. The successful knockdown of PRAME was confirmed by RT-qPCR and western blot (**Figure 10B**, **10C**). In consistent with the bioinformatics analysis results above, colony formation assays showed that PRAME knockdown notably inhibited the PC-3 and DU145 cell proliferation (**Figure 10D**). The wound healing assay and transwell assay respectively revealed that PRAME knockdown significantly suppressed the cell migration and invasion of PC-3 and DU145 (**Figure 11A**, **11B**), suggesting the growth promoting role of PRAME in PCa cells.

Discussion

Since PCa development is a slow process, it is important to distinguish patients with aggressive PCa from those with indolent PCa after the

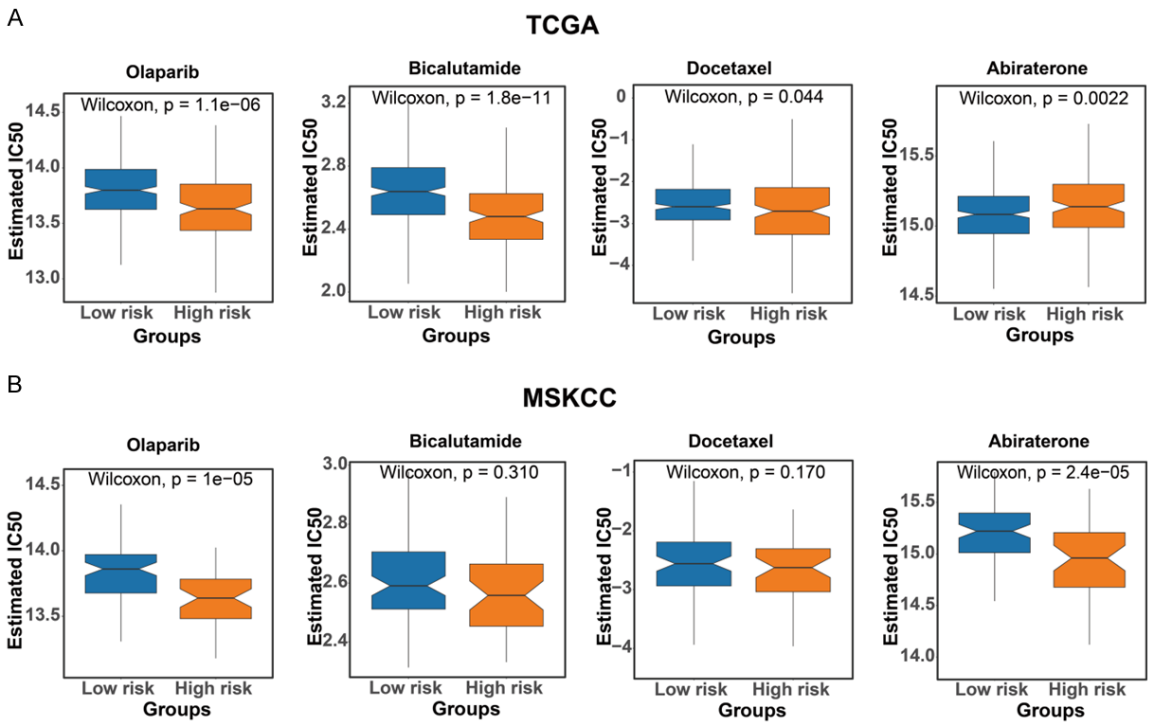


Figure 8. Drug sensitivity in low- and high-risk groups. The sensitivity to olaparib, bicalutamide, docetaxel and abiraterone in low- and high-risk groups in the (A) discovery and in the (B) validation cohorts.

initial therapy. BCR had been confirmed as an early sign of PCa progression and metastasis [7], thus, predicting the BCR of PCa is clinically helpful. In this study, we comprehensively examined the gene signature associated with BCR-free survival, and an 8-gene signature that could accurately predict BCR-free survival was established and validated in both the discovery and validation cohort (GSE21302).

Numerous studies have focused on genetic and epigenetic alterations as potential prognostic and predictive biomarkers for PCa in recent years [19]. Differential mRNA expression is involved in all biological processes such as cycle regulation, cell adhesion, angiogenesis, and tumorigenesis, which have been shown to play important roles in the tumorigenesis of PCa [20, 21], and therefore could serve as potential PCa prognostic indicators [22]. In this study, based on the mRNA expression profiles in TCGA PCa cohort, we identified 271 DEGs between PCa and paracancerous samples. Further univariate Cox analysis, LASSO and stepwise selection analysis stratified 8 genes (APOF, Clorf64, RPE65 and SEMG1, ARHGDIG, COMP, MKI67 and PRAME) to be significantly associated with the BCR-free survival of PCa, suggesting the potential application of these genes in prognosis prediction.

We then constructed a predictive signature based on these 8 genes and calculated the risk score for each sample. In modeling TCGA cohort, this 8-gene signature could robustly predict the BCR risk, which further verified in the validation cohort. Furthermore, the risk score was an independent risk factor for the BCR risk in both two cohorts, with a higher score representing an increased risk of BCR. Notable, the performance of this 8-gene signature was better than two previously reported signatures that consisted of 3 or 8 metabolic or pyroptosis-related genes [11, 12]. We noticed none of the two published signature genes were intersected with the 8 genes used in our study. As tumor stage and Gleason score were two crucial prognostic factors of PCa [23], we further combined them with the risk score to establish a nomogram and found this nomogram showed enhanced predict ability for BCR risk in both cohorts.

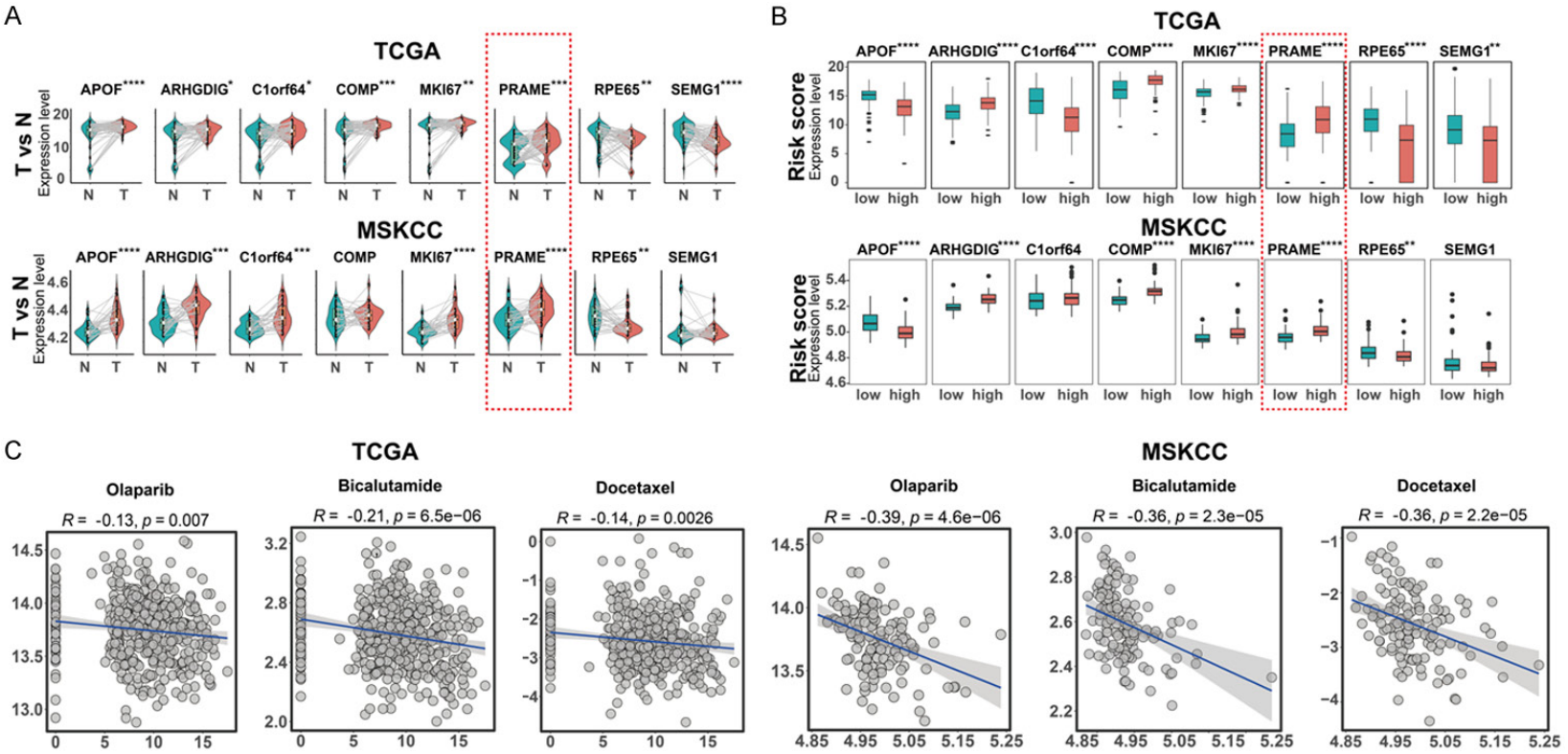


Figure 9. Gene expression in PCa tissues and in different risk groups and the correlation of PRAME expression with olaparib, docetaxel and bicalutamide responses in two cohorts. The expression of APOF, ARHGDIG, Clorf64, COMP, MKI67, PRAME, RPE65 and SEMG1 across (A) PCa tissues, (B) risk groups, and (C) the correlations of PRAME expression with olaparib, docetaxel and bicalutamide responses in the TCGA cohort and the MSKCC cohort (*P<0.05; **P<0.01; ***P<0.001; ****P<0.0001). T, tumor tissues; N, paracancerous normal tissues; PCa, prostate cancer.

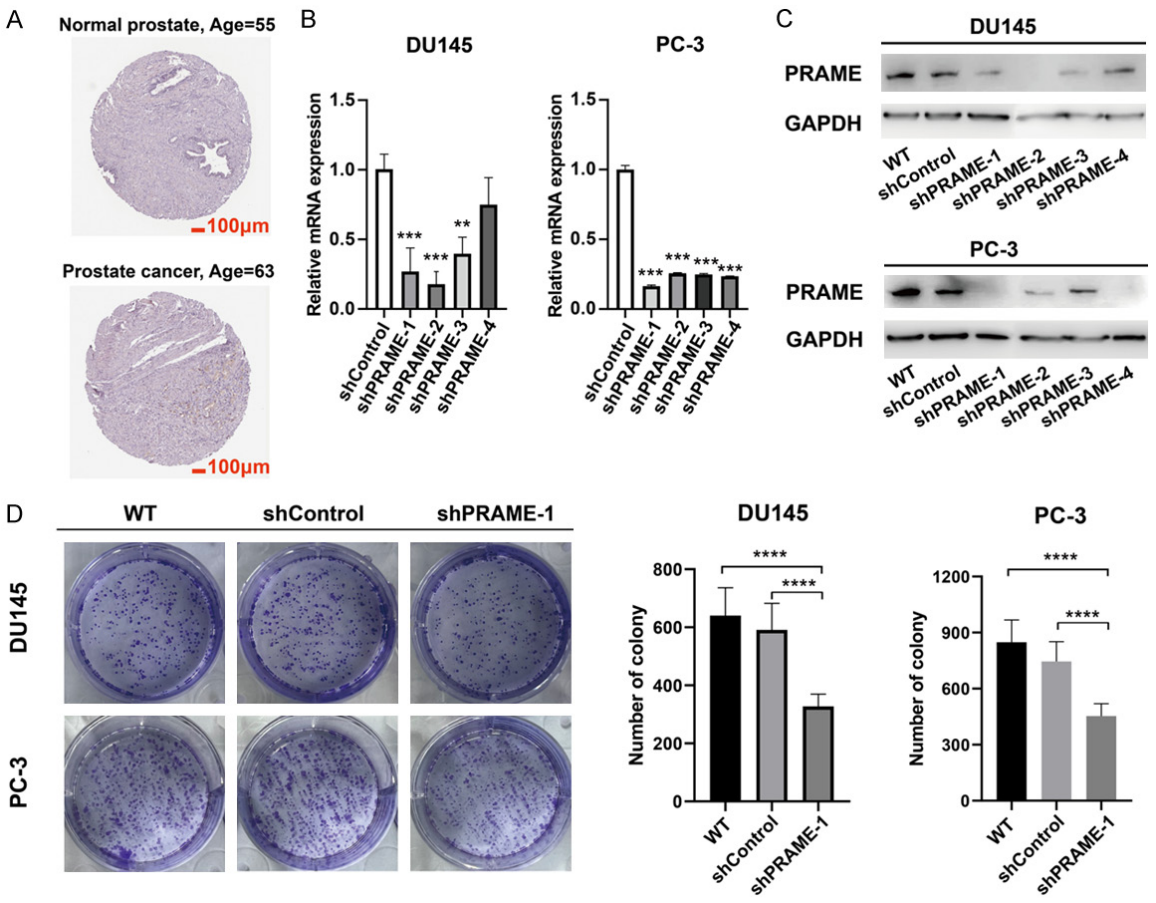


Figure 10. PRAME promote the proliferation of PCa cells. (A) IHC of the PRAME expression in PCa and normal tissues; (B) The efficiency of PRAME knockdown was indicated by RT-qPCR and (C) Western blot in DU145 and PC-3 cell lines; (D) Knockdown of PRAME inhibited the proliferation of DU145 and PC-3 cell, as determined by colony formation assay. PCa, prostate cancer.

Since drug therapies such as endocrine therapy and chemotherapy are also important for PCa treatment in addition to radical surgery and radiotherapy, we predicted the response of PCa patients to Abiraterone, docetaxel, olaparib and bicalutamide, the common drugs used clinical practice. We revealed that patients in the high-risk group responded better to olaparib than those in the low-risk group in both cohorts, suggesting the clinical importance of the 8-gene signature in optimizing treatment options for patients.

Among the 8 genes used in our signature, PRAME was not only showed the higher expression in the PCa tumor tissues and in the high-risk group, but also significantly related to drug response. Higher PRAME expression was also observed in the metastatic PCa and in most TCGA cancers. Together, these findings indicate the potential role of PRAME as a therapeutic target. It has been known that PRAME is Cancer/Testis genes (CT) expressed in germline cells but is activated in cancer cells that encode a tumor testis antigen which provides a growth advantage to cancer cells [24]. PRAME is the substrate recognition subunit of a Cullin2-RING (CRL2) E3 ubiquitin ligase and plays a role in chromatin regulation [25, 26]. However, the function of PRAME during PCa progression is limited as PRAME was mainly investigated in melanoma and hematologic malignancies. In this study, we demonstrated that PRAME knockdown inhibited the proliferation, migration, and invasion of PCa. Nonetheless, further studies are needed to determine how PRAME modulates PCa development.

As for the other genes in our 8-gene signature, COMP is reported to be a potent driver of PCa

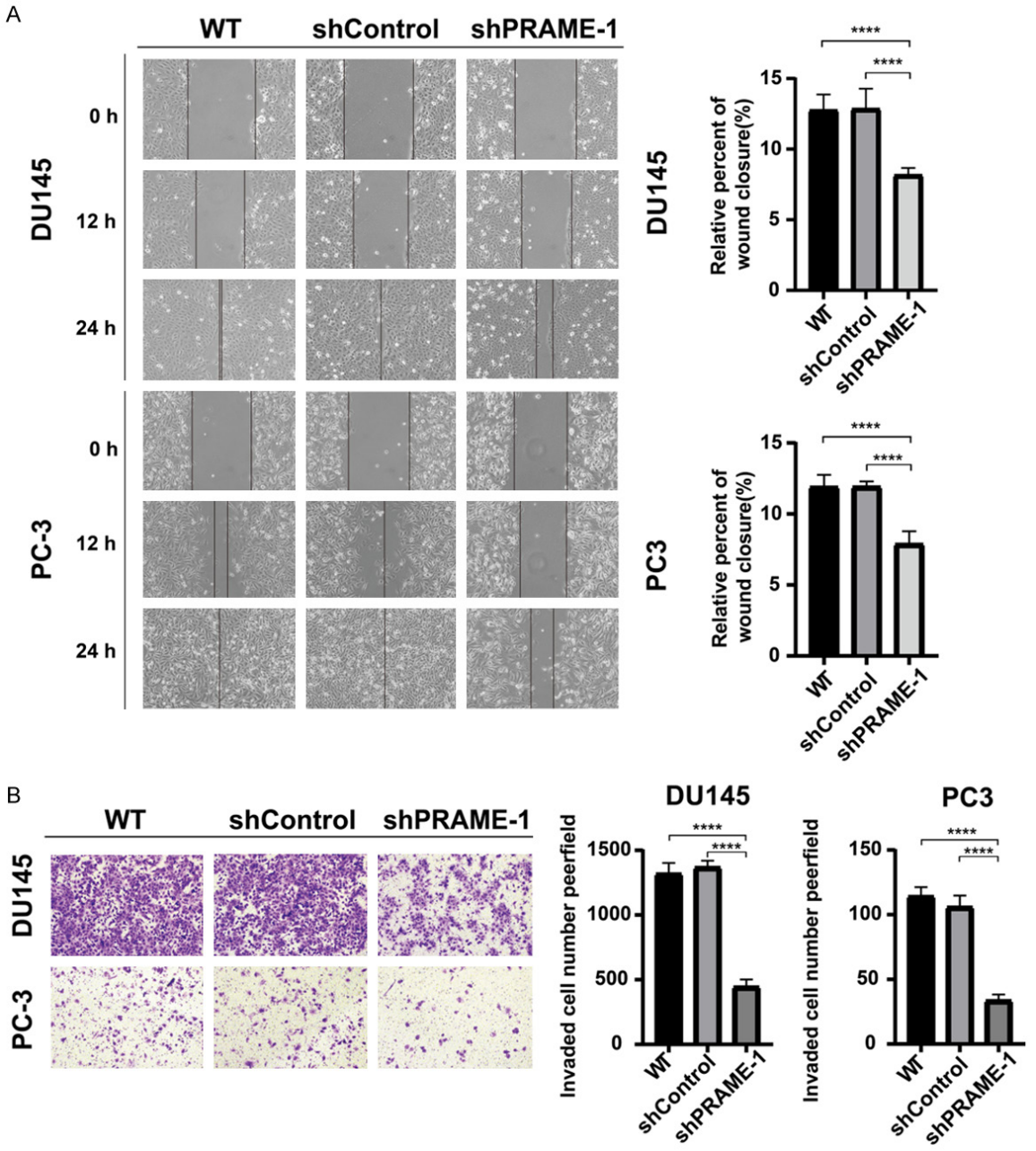


Figure 11. PRAME promote the cell migration and invasion in PCa. A. PRAME knockdown inhibited the migration of DU145 and PC-3 cells, as determined by wound healing assay; B. PRAME knockdown attenuated the invasion of DU145 and PC-3 cells, as determined by transwell assay. PCa, prostate cancer.

progression due to its anti-apoptotic effect via disturbing the Ca2+ homeostasis of cancer cells [27]. APOF is a prognostic factor for hepatocellular carcinoma [28], and its expression may be regulated by ETS-1/ETS-2 and C/EBPα [29]. C1orf64, also known as SRARP, may regulate the transcriptional function of androgen and estrogen receptors [30]. ARHGDIG is a GDP-dissociation inhibitor and is reported to be negatively associated with the malignancy of pancreatic cancer [31, 32]. Notable, MKI67, also called KI67, is proven as a robust index of cell cycle progression and cell proliferation [33]. Previous studies have also indicated its prognostic value in predicting BCR, progression and the overall survival of PCa [34], although

its high degree of variation among patients limits its clinical application. As for SEMG1, its expression seems to be both positively and negatively associated with PCa survival [35]. For example, one study reported that SEMG1 functioned as a co-activator of androgen receptor which prevents zinc-mediated cytotoxicity in PCa cells [36]. For RPE65, few studies report its downregulation in melanoma and lung cancer by analysing the public genome-wide expression profiles [37, 38]. In a recent pan-cancer analysis, RPE65 were found to be one of the gene signatures related to the loss of p53 function [39]. Collectively, these studies support the predictive role of the 8-gene signature in PCa BCR risk, but further studies are also needed to clarify the underlying mechanisms of these genes in PCa progression.

There are several limitations in this study. First, the 8-gene signature was constructed and evaluated based on retrospective public datasets which needs to be further verified in different genetic background PCa patients. Second, further mechanistic studies are needed to understand the role of these genes in PCa progression.

In conclusion, our study established an 8-gene signature which can accurately predict the BCR risk of PCa. Moreover, as a component of this 8-gene signature, PRAME enhanced the proliferation, migration, and invasion of PCa cells. These findings provide a novel tool for stratifying high-risk PCa patients and shed light on the mechanism of PCa progression.

Acknowledgements

This study is supported by grants from Shenzhen High-level Hospital Construction Fund, Sanming Project of Medicine in Shenzhen [No. SZSM202111003], Hospital Scientific Research Subject [SZ2020ZD003].

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Dongwen Wang, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, 113 Bao-he Road, Longgang District, Shenzhen 518116, China. E-mail: urology2007@126.com

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.

[2] Siegel RL, Miller KD, Fuchs HE and Jemal A. Cancer statistics, 2022. CA Cancer J Clin 2022; 72: 7-33.

[3] Xia C, Dong X, Li H, Cao M, Sun D, He S, Yang F, Yan X, Zhang S, Li N and Chen W. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. Chin Med J (Engl) 2022; 135: 584-590.

[4] Wang L, Lu B, He M, Wang Y, Wang Z and Du L. Prostate cancer incidence and mortality: global status and temporal trends in 89 countries from 2000 to 2019. Front Public Health 2022; 10: 811044.

[5] Mottet N, van den Bergh RCN, Briers E, Van den Broeck T, Cumberbatch MG, De Santis M, Fanti S, Fossati N, Gandaglia G, Gillessen S, Grivas N, Grummet J, Henry AM, van der Kwast TH, Lam TB, Lardas M, Liew M, Mason MD, Moris L, Oprea-Lager DE, van der Poel HG, Rouviere O, Schoots IG, Tilki D, Wiegel T, Willemse PM and Cornford P. EAU-EANM-ESTRO-ESUR-SIOG guidelines on prostate cancer-2020 update. Part 1: screening, diagnosis, and local treatment with curative intent. Eur Urol 2021; 79: 243-262.

[6] Suarez JF, Zamora V, Garin O, Gutierrez C, Pont A, Pardo Y, Goni A, Marino A, Hervas A, Herruzo I, Cabrera P, Sancho G, Ponce de Leon J, Macias V, Guedea F, Vigues F, Castells M and Ferrer M; Multicentric Spanish Group of Clinically Localized Prostate Cancer. Mortality and biochemical recurrence after surgery, brachytherapy, or external radiotherapy for localized prostate cancer: a 10-year follow-up cohort study. Sci Rep 2022; 12: 12589.

[7] Pound CR, Partin AW, Eisenberger MA, Chan DW, Pearson JD and Walsh PC. Natural history of progression after PSA elevation following radical prostatectomy. JAMA 1999; 281: 1591-1597.

[8] Gillessen S, Attard G, Beer TM, Beltran H, Bjartell A, Bossi A, Briganti A, Bristow RG, Chi KN, Clarke N, Davis ID, de Bono J, Drake CG, Duran I, Eeles R, Efstathiou E, Evans CP, Fanti S, Feng FY, Fizazi K, Frydenberg M, Gleave M, Halabi S, Heidenreich A, Heinrich D, Higano CTS, Hofman MS, Hussain M, James N, Kanesvaran R, Kantoff P, Khauli RB, Leibowitz R, Logothetis C,

Maluf F, Millman R, Morgans AK, Morris MJ, Mottet N, Mrabti H, Murphy DG, Murthy V, Oh WK, Ost P, O'Sullivan JM, Padhani AR, Parker C, Poon DMC, Pritchard CC, Reiter RE, Roach M, Rubin M, Ryan CJ, Saad F, Sade JP, Sartor O, Scher HI, Shore N, Small E, Smith M, Soule H, Sternberg CN, Steuber T, Suzuki H, Sweeney C, Sydes MR, Taplin ME, Tombal B, Turkeri L, van Oort I, Zapatero A and Omlin A. Management of patients with advanced prostate cancer: report of the Advanced Prostate Cancer Consensus Conference 2019. Eur Urol 2020; 77: 508-547.

[9] Williams IS, McVey A, Perera S, O'Brien JS, Kostos L, Chen K, Siva S, Azad AA, Murphy DG, Kasivisvanathan V, Lawrentschuk N and Frydenberg M. Modern paradigms for prostate cancer detection and management. Med J Aust 2022; 217: 424-433.

[10] Lin X, Kapoor A, Gu Y, Chow MJ, Xu H, Major P and Tang D. Assessment of biochemical recurrence of prostate cancer (review). Int J Oncol 2019; 55: 1194-1212.

[11] Zhao Y, Tao Z, Li L, Zheng J and Chen X. Predicting biochemical-recurrence-free survival using a three-metabolic-gene risk score model in prostate cancer patients. BMC Cancer 2022; 22: 239.

[12] Hu D, Cao Q, Tong M, Ji C, Li Z, Huang W, Jin Y, Tong G, Wang Y, Li P and Zhang H. A novel defined risk signature based on pyroptosis-related genes can predict the prognosis of prostate cancer. BMC Med Genomics 2022; 15: 24.

[13] Feng D, Shi X, Zhang F, Xiong Q, Wei Q and Yang L. Energy metabolism-related gene prognostic index predicts biochemical recurrence for patients with prostate cancer undergoing radical prostatectomy. Front Immunol 2022; 13: 839362.

[14] Zhang G and Luo Y. An immune-related lncRNA signature to predict the biochemical recurrence and immune landscape in prostate cancer. Int J Gen Med 2021; 14: 9031-9049.

[15] Lv Z, Wang J, Wang X, Mo M, Tang G, Xu H, Wang J, Li Y and Liu M. Identifying a ferroptosis-related gene signature for predicting biochemical recurrence of prostate cancer. Front Cell Dev Biol 2021; 9: 666025.

[16] Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME, Jones V, Bodycombe NE, Soule CK, Gould J, Alexander B, Li A, Montgomery P, Wawer MJ, Kuru N, Kotz JD, Hon CS, Munoz B, Liefeld T, Dancik V, Bittker JA, Palmer M, Bradner JE, Shamji AF, Clemons PA and Schreiber SL. Harnessing connectivity in a large-scale small-molecule sensitivity dataset. Cancer Discov 2015; 5: 1210-1223.

[17] Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Goncalves E, Barthorpe S, Lightfoot H, Cokelaer T, Greninger P, van Dyk E, Chang H, de Silva H, Heyn H, Deng X, Egan RK, Liu Q, Mironenko T, Mitropoulos X, Richardson L, Wang J, Zhang T, Moran S, Sayols S, Soleimani M, Tamborero D, Lopez-Bigas N, Ross-Macdonald P, Esteller M, Gray NS, Haber DA, Stratton MR, Benes CH, Wessels LFA, Saez-Rodriguez J, McDermott U and Garnett MJ. A landscape of pharmacogenomic interactions in cancer. Cell 2016; 166: 740-754.

[18] Maeser D, Gruener RF and Huang RS. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform 2021; 22: bbab260.

[19] Falahatkar R, Mokhtari G, Momeni-Moghaddam M, Teimoori M, Baghani Aal H, Akhavan A, Falahatkar S and Esmaeili S. An update on biochemical and genomic markers for prostate cancer. Urol J 2021; [Epub ahead of print].

[20] Fabbri L, Chakraborty A, Robert C and Vagner S. The plasticity of mRNA translation during cancer progression and therapy resistance. Nat Rev Cancer 2021; 21: 558-577.

[21] Massillo C, Dalton GN, Farre PL, De Luca P and De Siervi A. Implications of microRNA dysregulation in the development of prostate cancer. Reproduction 2017; 154: R81-R97.

[22] Dietrich D, Meller S, Uhl B, Ralla B, Stephan C, Jung K, Ellinger J and Kristiansen G. Nucleic acid-based tissue biomarkers of urologic malignancies. Crit Rev Clin Lab Sci 2014; 51: 173-199.

[23] D'Amico AV, Whittington R, Malkowicz SB, Wu YH, Chen M, Art M, Tomaszewski JE and Wein A. Combination of the preoperative PSA level, biopsy gleason score, percentage of positive biopsies, and MRI T-stage to predict early PSA failure in men with clinically localized prostate cancer. Urology 2000; 55: 572-577.

[24] Hofmann O, Caballero OL, Stevenson BJ, Chen YT, Cohen T, Chua R, Maher CA, Panji S, Schaefer U, Kruger A, Lehvaslaiho M, Carninci P, Hayashizaki Y, Jongeneel CV, Simpson AJ, Old LJ and Hide W. Genome-wide analysis of cancer/testis gene expression. Proc Natl Acad Sci U S A 2008; 105: 20422-20427.

[25] Costessi A, Mahrour N, Tijchon E, Stunnenberg R, Stoel MA, Jansen PW, Sela D, Martin-Brown S, Washburn MP, Florens L, Conaway JW, Conaway RC and Stunnenberg HG. The tumour antigen PRAME is a subunit of a Cul2 ubiquitin ligase and associates with active NFY promoters. EMBO J 2011; 30: 3786-3798.

[26] Lin HC, Ho SC, Chen YY, Khoo KH, Hsu PH and Yen HC. SELENOPROTEINS. CRL2 aids elimination of truncated selenoproteins produced by

failed UGA/Sec decoding. Science 2015; 349: 91-95.
[27] Englund E, Canesin G, Papadakos KS, Vishnu N, Persson E, Reitsma B, Anand A, Jacobsson L, Helczynski L, Mulder H, Bjartell A and Blom AM. Cartilage oligomeric matrix protein promotes prostate cancer progression by enhancing invasion and disrupting intracellular calcium homeostasis. Oncotarget 2017; 8: 98298-98311.
[28] Wang YB, Zhou BX, Ling YB, Xiong ZY, Li RX, Zhong YS, Xu MX, Lu Y, Liang H, Chen GH, Yao ZC and Deng MH. Decreased expression of ApoF associates with poor prognosis in human hepatocellular carcinoma. Gastroenterol Rep (Oxf) 2019; 7: 354-360.
[29] Shen XB, Huang L, Zhang SH, Wang DP, Wu YL, Chen WN, Xu SH and Lin X. Transcriptional regulation of the apolipoprotein F (ApoF) gene by ETS and C/EBPalpha in hepatoma cells. Biochimie 2015; 112: 1-9.
[30] Naderi A. C1orf64 is a novel androgen receptor target gene and coregulator that interacts with 14-3-3 protein in breast cancer. Oncotarget 2017; 8: 57907-57933.
[31] Adra CN, Iyengar AR, Syed FA, Kanaan IN, Rilo HL, Yu W, Kheraj R, Lin SR, Horiuchi T, Khan S, Weremowicz S, Lim B, Morton CC and Higgs DR. Human ARHGDIG, a GDP-dissociation inhibitor for Rho proteins: genomic structure, sequence, expression analysis, and mapping to chromosome 16p13.3. Genomics 1998; 53: 104-109.
[32] de Leon-Bautista MP, Cardenas-Aguayo MD, Casique-Aguirre D, Almaraz-Salinas M, Parraguirre-Martinez S, Olivo-Diaz A, Thompson-Bonilla MD and Vargas M. Immunological and functional characterization of RhoGDI3 and its molecular targets RhoG and RhoB in human pancreatic cancerous and normal cells. PLoS One 2016; 11: e0166370.
[33] Gerdes J, Schwab U, Lemke H and Stein H. Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. Int J Cancer 1983; 31: 13-20.
[34] Berlin A, Castro-Mesta JF, Rodriguez-Romo L, Hernandez-Barajas D, Gonzalez-Guerrero JF, Rodriguez-Fernandez IA, Gonzalez-Conchas G, Verdines-Perez A and Vera-Badillo FE. Prognostic role of Ki-67 score in localized prostate cancer: a systematic review and meta-analysis. Urol Oncol 2017; 35: 499-506.
[35] Shuvalov O, Kizenko A, Petukhov A, Fedorova O, Daks A and Barlev N. Emerging roles of cancer-testis antigenes, semenogelin 1 and 2, in neoplastic cells. Cell Death Discov 2021; 7: 97.
[36] Ishiguro H, Izumi K, Kashiwagi E, Zheng Y, Li Y, Kawahara T and Miyamoto H. Semenogelin I promotes prostate cancer cell growth via functioning as an androgen receptor coactivator and protecting against zinc cytotoxicity. Am J Cancer Res 2015; 5: 738-747.
[37] Hassel JC, Amann PM, Schadendorf D, Eichmuller SB, Nagler M and Bazhin AV. Lecithin retinol acyltransferase as a potential prognostic marker for malignant melanoma. Exp Dermatol 2013; 22: 757-759.
[38] Zhou D, Sun Y, Jia Y, Liu D, Wang J, Chen X, Zhang Y and Ma X. Bioinformatics and functional analyses of key genes in smoking-associated lung adenocarcinoma. Oncol Lett 2019; 18: 3613-3622.
[39] Keshavarz-Rahaghi F, Pleasance E, Kolisnik T and Jones SJM. A p53 transcriptional signature in primary and metastatic cancers derived using machine learning. Front Genet 2022; 13: 987238.

Table S1. General characteristics of patients with paired tumor and normal specimens in TCGA cohort

	Overall (n=52)
Age, years	
Median (Min, Max)	61.0 (43.0, 72.0)
Gleason score	
6	5 (9.6)
7	40 (76.9)
8	3 (5.8)
9	4 (7.7)
PSA level (ng/ml)	
Median (Min, Max)	0.1 (0, 11.2)
Tumor stage	
T2	29 (55.8)
T3	21 (40.4)
T4	2 (3.8)
Lymph node status	
N0	46 (88.5)
N1	1 (1.9)
Not reported	5 (9.6)
Vital status	
Alive	52 (100.0)
Dead	0 (0)
BCR status	
No	41 (78.8)
Yes	2 (3.8)
Not reported	9 (17.3)
Median follow-up time (Months)	6.3

BCR, biochemical recurrence; TCGA, The Cancer Genome Atlas; PSA, prostate-specific antigen.

Table S2. General characteristics of patients with complete BCR information in TCGA and MSKCC cohorts

	TCGA (n=446)	MSKCC (n=131)
Age, years		
Median (Min, Max)	61.0 (43.0, 77.0)	58.0 (37.3, 83.0)
Race		
Asian	12 (2.7)	2 (1.5)
White	365 (81.8)	98 (74.8)
Black	55 (12.4)	27 (20.6)
Not reported	14 (3.1)	4 (3.1)
Gleason score		
6	41 (9.2)	41 (31.3)
7	225 (50.4)	74 (56.5)
8	54 (12.1)	8 (6.1)
9	123 (27.6)	7 (5.3)
10	3 (0.7)	0 (0)
Not reported	0 (0)	1 (0.8)
PSA level (ng/ml)		
Median (Min, Max)	0.1 (0, 39.8)	5.8 (1.09, 132.0)
Pathological tumor stage		
T2	176 (39.5)	85 (64.9)
T3	256 (57.4)	40 (30.5)
T4	9 (2.0)	6 (4.6)
Not reported	5 (1.1)	0 (0)
Lymph node status		
N0	312 (70.0)	102 (77.9)
N1	67 (15.0)	6 (4.6)
Not reported	67 (15.0)	23 (17.6)
Vital status		
Alive	444 (99.6)	124 (94.7)
Dead	2 (0.4)	7 (5.3)
BCR status		
No	383 (85.9)	104 (79.4)
Yes	63 (14.1)	27 (20.6)
Median follow-up time (Months)	23.5	46.5

BCR, biochemical recurrence; MSKCC, MSK-IMPACT Clinical Sequencing Cohort; PSA, prostate-specific antigen; TCGA, The Cancer Genome Atlas.

Table S3. Differentially expressed genes between tumor group and paired normal sample group

Gene name	Tumor (n=52) *v.s.* Normal (n=52)		
	logFC[a]	*P* value	*P* adjust value[c] (FDR)
Up regulated in tumor compared to normal[b]			
DLX1	6.44	3.28E-82	5.56E-78
ZIC2	5.71	5.94E-70	5.03E-66
NKX2-3	5.14	3.28E-54	8.57E-51
DLX2	5.04	8.87E-58	3.76E-54
ADAM2	4.56	3.72E-31	4.23E-29
TDRD1	4.50	7.68E-48	9.29E-45
SLC45A2	4.23	3.54E-54	8.57E-51
UGT2B4	3.81	7.99E-29	6.09E-27
OR51E2	3.69	7.14E-28	4.65E-26
PHGR1	3.66	6.51E-38	1.72E-35
MMP26	3.61	1.62E-38	4.73E-36
HOXC6	3.58	2.36E-47	2.35E-44
ANKRD34B	3.41	5.61E-17	7.28E-16
B3GNT6	3.36	3.77E-17	5.03E-16
SIM2	3.31	1.08E-52	2.28E-49
AMACR	3.29	2.84E-62	1.60E-58
SERPINA11	3.28	3.96E-20	7.74E-19
GAL	3.26	1.55E-41	6.74E-39
APOF	3.25	3.12E-34	5.28E-32
FOXD1	3.22	6.07E-24	2.13E-22
MNX1	3.21	1.77E-20	3.64E-19
COL10A1	3.14	3.13E-27	1.84E-25
EPHA8	3.12	1.12E-23	3.80E-22
ONECUT2	3.09	4.56E-43	2.15E-40
PAX1	3.08	4.63E-16	5.30E-15
SLIT1	3.07	4.57E-36	9.79E-34
TGM3	2.95	2.27E-35	4.58E-33
ADAM7	2.94	5.52E-09	2.51E-08
FGL1	2.93	5.08E-21	1.14E-19
NLRP12	2.91	3.83E-51	6.48E-48
EN2	2.86	1.47E-14	1.37E-13
ACSM1	2.82	5.84E-44	3.09E-41
DNAH5	2.81	9.03E-45	5.27E-42
KRT20	2.80	2.09E-12	1.48E-11
OTX1	2.77	2.85E-18	4.37E-17
TPO	2.76	2.15E-25	9.55E-24
HPN	2.74	1.24E-52	2.33E-49
LMX1B	2.73	2.48E-31	2.86E-29
TERT	2.73	2.14E-16	2.56E-15
HOXC4	2.69	6.49E-28	4.28E-26
CCDC83	2.67	9.36E-21	2.01E-19
ATP8A2	2.66	4.67E-20	9.03E-19
PRAME	2.64	1.28E-09	6.36E-09
GLYATL1	2.60	3.54E-21	8.17E-20
C2orf72	2.58	9.17E-29	6.81E-27
LUZP2	2.57	1.68E-20	3.47E-19

ARHGAP19-SLIT1	2.56	5.15E-16	5.83E-15
KLK15	2.55	1.55E-23	5.09E-22
B3GAT1	2.53	1.36E-27	8.51E-26
CST2	2.52	7.24E-19	1.20E-17
UNC5A	2.50	7.17E-16	7.99E-15
BEND4	2.49	2.52E-20	5.09E-19
HJURP	2.47	8.92E-42	3.97E-39
CGREF1	2.46	1.24E-45	7.79E-43
UGT1A3	2.45	3.50E-10	1.87E-09
FEV	2.43	1.23E-19	2.24E-18
KCNG3	2.40	3.92E-20	7.68E-19
INSM1	2.38	1.99E-15	2.07E-14
FFAR2	2.37	1.06E-23	3.64E-22
EBF2	2.36	5.22E-30	4.80E-28
GHRHR	2.36	5.91E-23	1.78E-21
RPRML	2.34	2.09E-11	1.31E-10
ANKRD66	2.33	1.54E-17	2.15E-16
GNG13	2.33	5.23E-16	5.93E-15
ZNF560	2.31	1.16E-08	5.03E-08
CCL18	2.30	1.92E-15	2.01E-14
IGSF23	2.30	2.89E-10	1.56E-09
PDLIM5	2.28	2.55E-47	2.40E-44
CHIT1	2.28	3.63E-15	3.64E-14
MATK	2.27	1.90E-36	4.46E-34
COMP	2.26	4.32E-19	7.33E-18
SHISA8	2.25	1.82E-14	1.67E-13
SLCO1A2	2.25	1.89E-10	1.04E-09
ABCC4	2.23	5.24E-27	2.99E-25
GOLM1	2.23	8.19E-34	1.28E-31
FOLH1	2.23	2.49E-16	2.94E-15
PPM1E	2.22	2.08E-26	1.08E-24
HIST1H2AI	2.22	6.99E-18	1.01E-16
NETO2	2.21	9.70E-47	8.21E-44
CACNA1D	2.21	6.66E-34	1.05E-31
TBX10	2.20	8.26E-12	5.44E-11
FRMPD3	2.19	4.79E-29	3.81E-27
MYBL2	2.17	7.66E-33	1.06E-30
RNF157	2.16	7.85E-25	3.20E-23
ELAVL2	2.16	3.95E-17	5.25E-16
NLRP8	2.16	1.28E-19	2.31E-18
NKX2-2	2.15	7.64E-08	2.99E-07
ANGPTL3	2.13	1.76E-13	1.43E-12
TMEM178A	2.13	8.48E-16	9.33E-15
MKI67	2.12	3.79E-29	3.10E-27
MMP10	2.12	4.49E-12	3.05E-11
SMPDL3B	2.12	9.68E-41	3.73E-38
CTD-2501B8.1	2.12	5.07E-10	2.65E-09
SLCO1B3	2.10	9.80E-08	3.77E-07
NPY	2.09	2.31E-08	9.63E-08
NUTM2F	2.09	1.93E-10	1.06E-09

EPHA10	2.09	1.06E-47	1.20E-44
ANXA10	2.09	5.35E-14	4.65E-13
PTPRT	2.09	7.24E-14	6.19E-13
RRM2	2.07	3.40E-29	2.81E-27
TUBB4A	2.06	1.41E-14	1.32E-13
C1orf64	2.05	1.02E-10	5.84E-10
DCSTAMP	2.05	6.56E-11	3.86E-10
TRPC7	2.04	4.18E-08	1.69E-07
NEIL3	2.04	1.61E-21	3.91E-20
ARHGDIG	2.04	2.24E-17	3.06E-16
TFF3	2.04	2.02E-10	1.11E-09
HAO1	2.02	5.68E-08	2.26E-07
COL2A1	2.01	6.58E-07	2.25E-06
TECTB	2.01	5.72E-07	1.97E-06
Down regulated in tumor compared to normal[b]			
AQP2	-4.47	1.98E-18	3.10E-17
PADI3	-3.83	2.51E-23	7.92E-22
SLC39A2	-3.78	6.18E-26	2.97E-24
MYH6	-3.67	3.92E-29	3.19E-27
PNMT	-3.49	1.61E-30	1.68E-28
C10orf99	-3.46	3.01E-21	7.00E-20
SEMG1	-3.30	5.59E-11	3.32E-10
KRT24	-3.26	1.94E-14	1.78E-13
BMP5	-3.24	4.12E-14	3.61E-13
PIK3C2G	-3.24	1.21E-34	2.21E-32
FOXI1	-3.20	2.32E-26	1.18E-24
HS3ST5	-3.18	3.49E-16	4.07E-15
LRRTM3	-3.16	8.74E-18	1.25E-16
KRT13	-3.14	4.71E-18	7.04E-17
GPRC6A	-3.12	4.29E-18	6.45E-17
DMRT2	-3.06	3.86E-22	1.03E-20
NPFFR2	-3.06	1.58E-14	1.47E-13
UGT3A1	-3.03	9.15E-15	8.76E-14
TNNT3	-2.98	7.80E-20	1.46E-18
HSPB3	-2.95	1.44E-16	1.78E-15
KCNH5	-2.91	1.23E-18	1.99E-17
ACTC1	-2.91	9.70E-35	1.81E-32
SERPINA5	-2.88	3.21E-19	5.50E-18
EMX2	-2.87	2.05E-16	2.47E-15
SCGB3A1	-2.86	1.30E-22	3.73E-21
LINGO2	-2.84	1.10E-21	2.71E-20
SCGB1A1	-2.83	2.81E-13	2.21E-12
DCC	-2.83	1.75E-30	1.80E-28
COLEC10	-2.77	1.67E-14	1.54E-13
FABP7	-2.75	3.07E-18	4.70E-17
SLCO4C1	-2.73	5.55E-16	6.26E-15
DAPL1	-2.73	2.07E-18	3.23E-17
CHRM2	-2.73	1.69E-10	9.39E-10
CLCA2	-2.71	1.78E-23	5.80E-22
MYH2	-2.71	1.62E-08	6.93E-08

GPX2	-2.68	2.25E-29	1.94E-27
ADAMTS18	-2.67	4.21E-18	6.33E-17
FAM83C	-2.64	1.61E-12	1.15E-11
CA3	-2.62	1.07E-14	1.01E-13
CBLN4	-2.61	2.35E-16	2.80E-15
MCF2	-2.60	1.86E-23	6.02E-22
SFRP5	-2.60	3.24E-18	4.94E-17
NPPC	-2.60	8.67E-19	1.43E-17
DUSP13	-2.60	1.33E-09	6.58E-09
KCNJ15	-2.59	2.90E-28	2.00E-26
ATP6V0A4	-2.58	4.73E-20	9.12E-19
PENK	-2.56	1.02E-31	1.26E-29
LGR6	-2.54	2.73E-35	5.44E-33
C10orf82	-2.51	3.93E-27	2.29E-25
AKR1B15	-2.51	1.88E-11	1.18E-10
TMEM213	-2.50	1.13E-16	1.40E-15
KRT222	-2.50	2.21E-30	2.23E-28
PLA2G3	-2.49	2.47E-21	5.86E-20
PI16	-2.49	1.08E-19	2.00E-18
GABRA3	-2.47	4.21E-20	8.20E-19
FUT3	-2.46	6.94E-18	1.01E-16
CDH8	-2.46	6.68E-24	2.34E-22
CRTAC1	-2.45	7.25E-22	1.85E-20
MORC1	-2.45	2.52E-19	4.39E-18
LY6D	-2.45	5.36E-12	3.60E-11
CYP4F22	-2.45	1.31E-21	3.21E-20
KRT9	-2.42	5.58E-20	1.07E-18
CNTNAP4	-2.42	8.05E-09	3.59E-08
SBSPON	-2.42	3.23E-29	2.68E-27
CIDEC	-2.41	2.50E-24	9.35E-23
MSLN	-2.40	2.12E-18	3.31E-17
S100A14	-2.40	2.91E-29	2.44E-27
FAM163A	-2.40	5.05E-14	4.39E-13
ZSCAN4	-2.40	2.65E-15	2.72E-14
HSD17B13	-2.39	3.35E-13	2.62E-12
GLRA4	-2.38	2.05E-26	1.06E-24
CAPNS2	-2.37	5.50E-15	5.43E-14
PON3	-2.37	2.03E-25	9.02E-24
PDE1C	-2.37	1.51E-29	1.32E-27
CHP2	-2.36	6.03E-23	1.81E-21
RP11-569G13.3	-2.35	1.20E-19	2.20E-18
SEMG2	-2.34	8.68E-06	2.51E-05
C11orf87	-2.33	9.05E-12	5.92E-11
KCNJ3	-2.33	8.98E-21	1.94E-19
MAL	-2.33	7.93E-36	1.68E-33
LY6G6D	-2.32	1.66E-21	4.02E-20
CKM	-2.31	5.39E-17	7.02E-16
UGT2B7	-2.31	4.42E-07	1.55E-06
BSND	-2.30	8.22E-15	7.93E-14
ADRB3	-2.30	3.24E-09	1.51E-08

ACTA1	-2.29	3.56E-14	3.14E-13
FOLR1	-2.29	6.05E-27	3.39E-25
C14orf180	-2.29	4.73E-13	3.63E-12
CPNE6	-2.29	1.43E-27	8.93E-26
CXCL13	-2.28	9.23E-14	7.74E-13
TNMD	-2.27	7.70E-14	6.54E-13
ODAM	-2.27	3.17E-26	1.59E-24
NHLH2	-2.26	1.89E-15	1.98E-14
EVX2	-2.26	1.75E-22	4.92E-21
CYP4B1	-2.26	8.02E-29	6.09E-27
FAM163B	-2.26	3.53E-10	1.88E-09
KRT16	-2.25	5.97E-22	1.55E-20
CCNI2	-2.24	4.33E-18	6.49E-17
KRT4	-2.24	1.69E-09	8.23E-09
ARSF	-2.24	1.95E-22	5.44E-21
HRASLS	-2.22	1.40E-07	5.26E-07
LGALS7B	-2.22	8.78E-16	9.63E-15
PCP4L1	-2.21	3.62E-34	5.94E-32
WIF1	-2.21	3.48E-15	3.50E-14
SYT8	-2.20	1.07E-15	1.16E-14
DUOXA1	-2.19	1.12E-26	6.06E-25
SLC9A4	-2.19	7.17E-11	4.20E-10
KCNF1	-2.18	3.23E-21	7.45E-20
CA14	-2.18	4.55E-31	5.14E-29
MEI4	-2.17	1.16E-11	7.50E-11
C2orf88	-2.17	5.05E-41	2.08E-38
SERPINB11	-2.17	1.71E-07	6.37E-07
P2RX6	-2.17	1.47E-34	2.62E-32
LRRC3B	-2.16	3.80E-13	2.94E-12
SNCG	-2.16	6.84E-46	4.45E-43
FAM83A	-2.15	4.15E-11	2.51E-10
QPRT	-2.15	5.98E-40	2.15E-37
HPR	-2.15	1.49E-10	8.34E-10
AQP5	-2.14	1.14E-22	3.33E-21
CYP11A1	-2.14	4.62E-46	3.13E-43
SCARA5	-2.13	2.76E-18	4.24E-17
MYH7	-2.12	2.09E-05	5.66E-05
KY	-2.12	3.19E-30	3.07E-28
WISP2	-2.12	1.88E-16	2.28E-15
ASPA	-2.11	1.69E-33	2.57E-31
DPT	-2.10	1.28E-23	4.29E-22
SPINK2	-2.10	6.62E-08	2.61E-07
GSTP1	-2.09	6.67E-47	5.94E-44
STAC	-2.08	6.04E-21	1.33E-19
CRABP2	-2.08	5.08E-29	3.99E-27
GJB5	-2.08	1.08E-20	2.30E-19
MPZ	-2.07	3.02E-24	1.11E-22
FRMD7	-2.07	6.68E-14	5.74E-13
TAF7L	-2.06	6.67E-25	2.76E-23
CHRNA4	-2.06	7.72E-11	4.50E-10

C1QL1	-2.06	6.08E-18	8.92E-17
SLC46A2	-2.06	7.28E-11	4.26E-10
IGSF1	-2.05	5.14E-25	2.15E-23
RBFOX1	-2.04	6.21E-23	1.85E-21
CNTFR	-2.04	3.96E-27	2.31E-25
CXCR2	-2.04	9.13E-29	6.81E-27
APOBEC2	-2.04	5.20E-11	3.11E-10
BRINP2	-2.04	2.19E-09	1.05E-08
CCK	-2.04	1.86E-08	7.88E-08
C2orf71	-2.04	1.23E-15	1.32E-14
RPE65	-2.03	5.72E-18	8.45E-17
CYP3A5	-2.03	5.00E-29	3.96E-27
TRH	-2.03	5.35E-07	1.85E-06
AOX1	-2.03	4.55E-43	2.15E-40
SOSTDC1	-2.03	2.30E-28	1.62E-26
HPCAL4	-2.02	2.18E-14	1.97E-13
SERPINB5	-2.02	1.46E-14	1.36E-13
EPHB1	-2.01	8.71E-31	9.52E-29
RHCG	-2.01	6.78E-12	4.51E-11
VSNL1	-2.01	7.29E-15	7.08E-14
PTGS1	-2.01	1.32E-18	2.11E-17
ALDH3A1	-2.01	1.80E-31	2.14E-29
SLC30A8	-2.01	4.53E-07	1.58E-06
OPTC	-2.01	7.63E-13	5.67E-12
TMEM40	-2.01	8.29E-23	2.44E-21
COL17A1	-2.00	3.81E-20	7.48E-19

[a]The fold change (FC) was calculated as the ratio of gene expression in tumor tissues to the paired normal tissues. [b]The significantly altered genes (*P* adjust value <0.05) in tumor tissue compared to paired normal tissue with |logFC| >2 were listed. [c]DESeq2 paired analysis and P adjust values were used after adjustment for multiple comparison by FDR.

Table S4. Identification of the eight genes and their association with BCR-free survival in PCa

	HR (95% CI)	*P* value
APOF	0.82 (0.14-0.91)	<0.001
ARHGDIG	1.22 (1.06-1.40)	0.006
C1orf64	0.90 (0.85-0.96)	<0.001
COMP	1.26 (1.08-1.47)	0.003
MKI67	1.41 (1.10-1.81)	0.007
PRAME	1.10 (1.03-1.17)	0.006
RPE65	0.91 (0.87-0.96)	<0.001
SEMG1	0.95 (0.90-0.99)	0.049

BCR, biochemical recurrence; CI, confidence interval; HR, hazard ratio; PCa, prostate cancer.

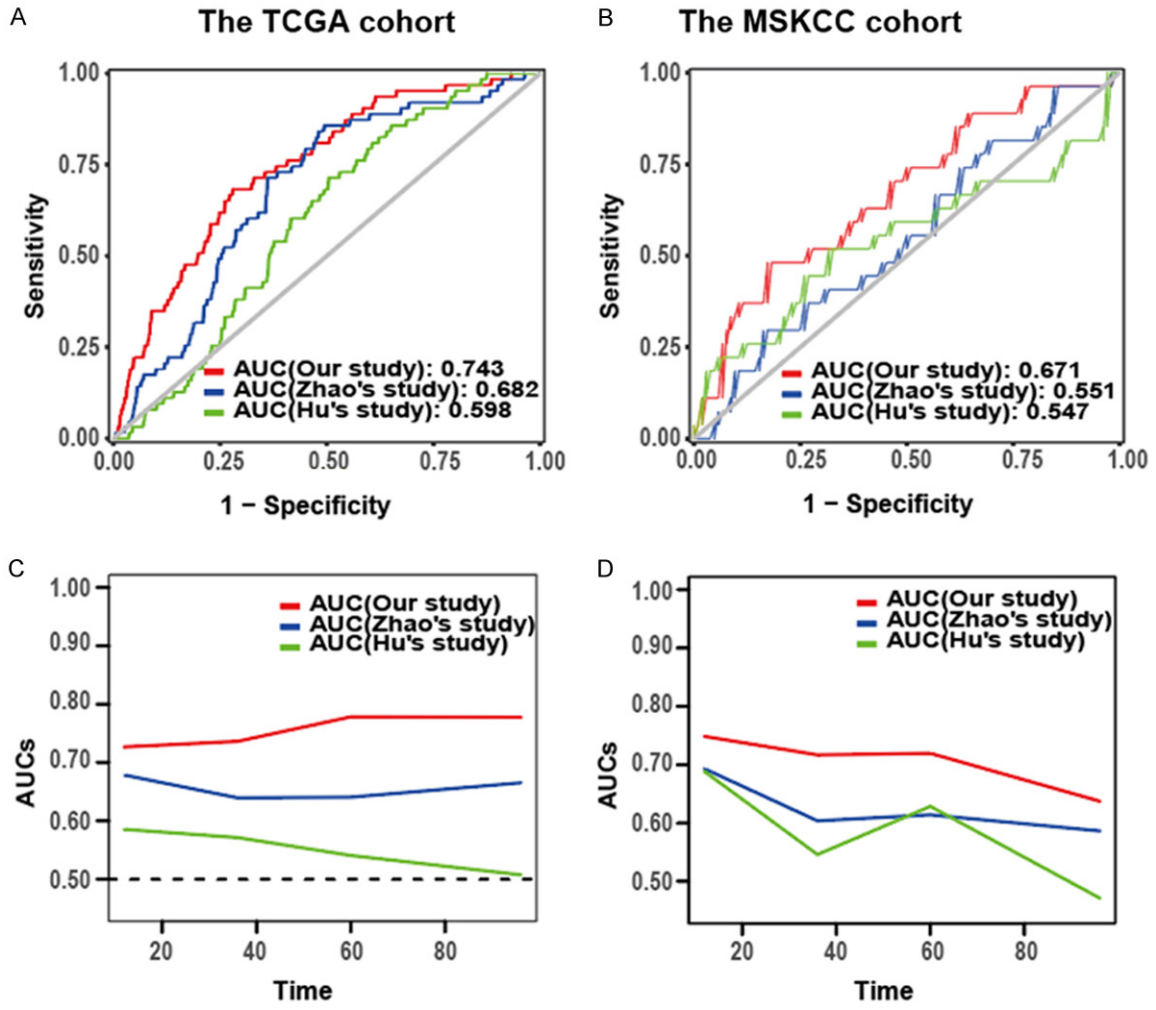


Figure S1. Performance of the 8-gene signature and reported signatures in different cohorts. The area under the receiver operating characteristic (ROC) curves for the BCR discrimination ability in the TCGA cohort (A) and the MSKCC cohort (B). The AUCs of the 8-gene signature and reported signatures at 1-, 3-, 5-, and 8-year in the TCGA cohort (C) and the MSKCC cohort (D).

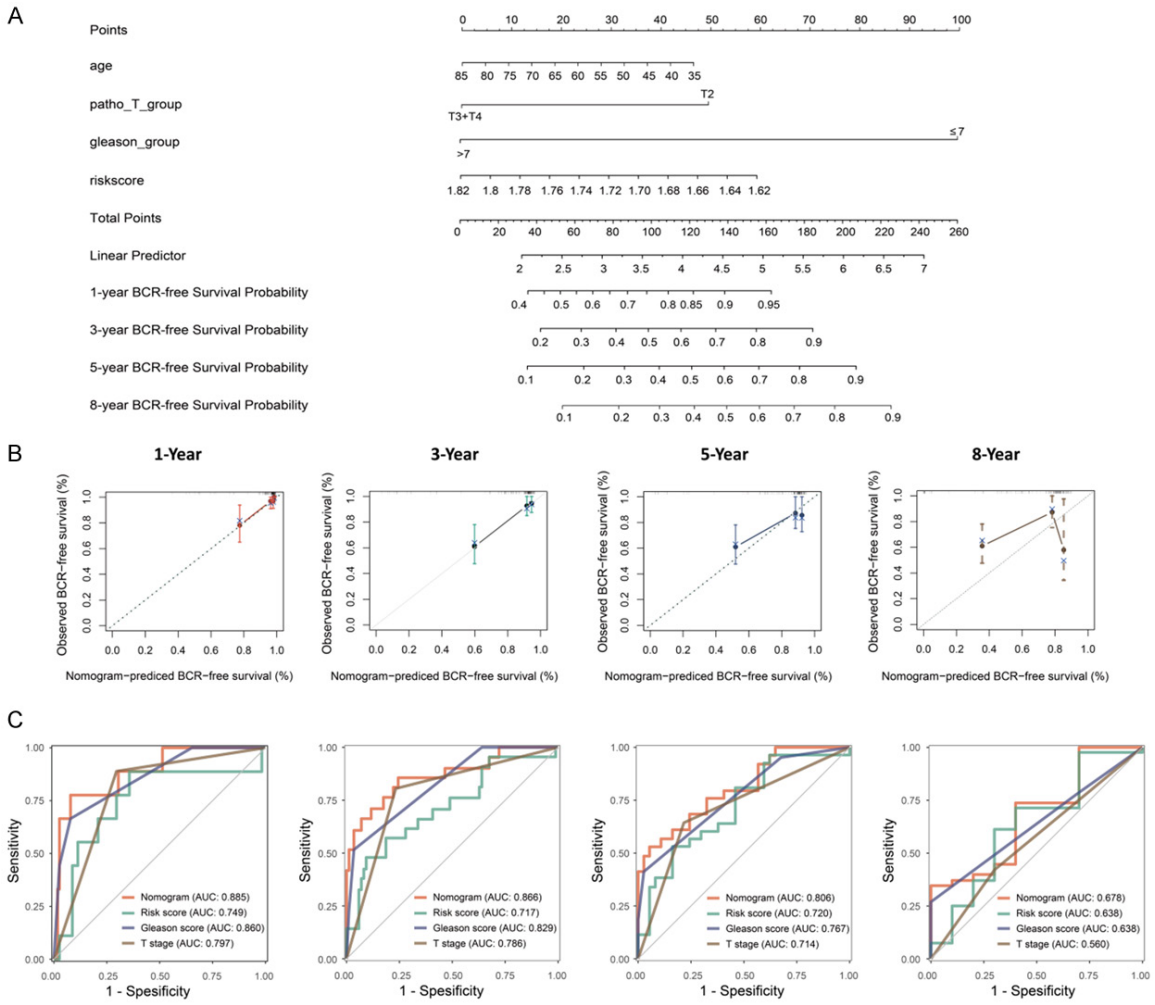
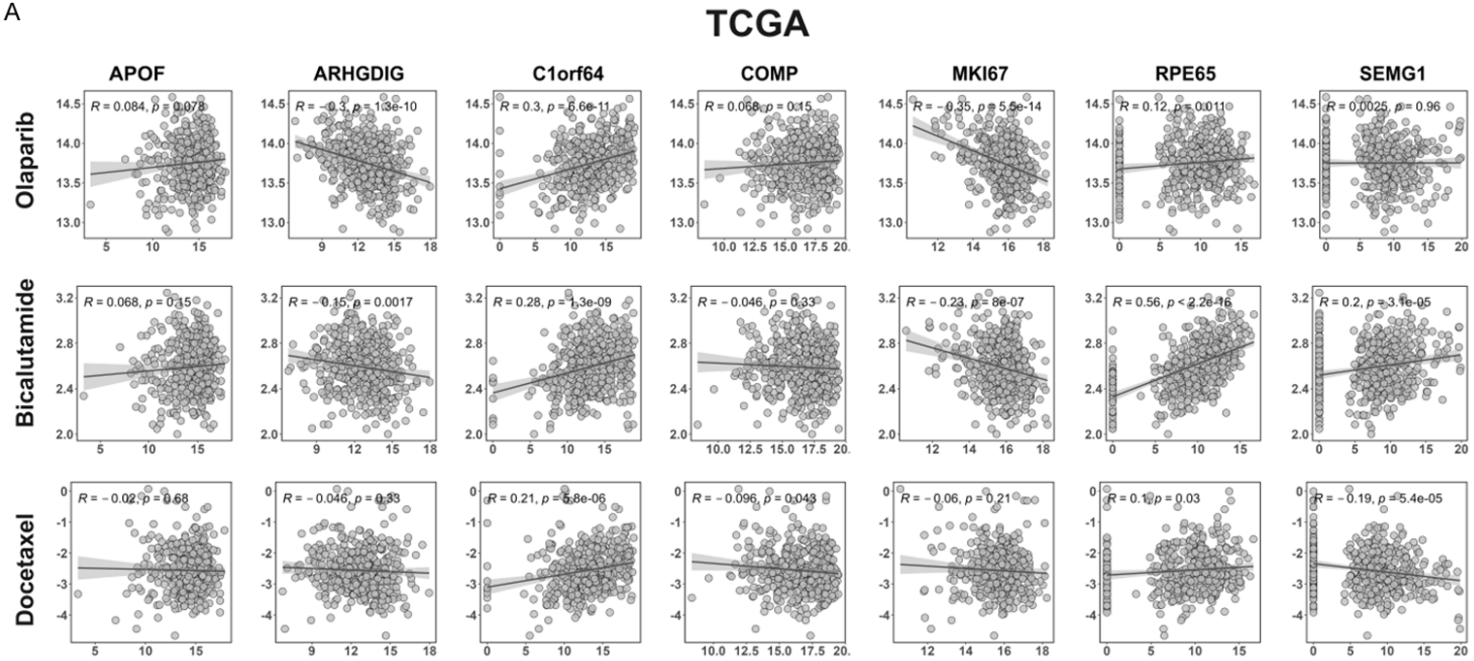


Figure S2. Nomogram for predicting the BCR-free survival of PCa patients in the MSKCC cohort. A. A prognostic nomogram including signature risk score and clinical factors. B. The calibration curves of the 1-, 3-, 5-, and 8-year BCR-free survival. C. ROC curve used to evaluate the predictive performance of 1-, 3-, 5-, and 8-year BCR-free survival. BCR, biochemical recurrence; PCa, prostate cancer.

A


TCGA
APOF
ARHGDIG
C1orf64
COMP
MKI67
RPE65
SEMG1
Olaparib
Bicalutamide
Docetaxel
R = 0.084, p = 0.078
R = −0.3, p = 1.3e-10
R = 0.3, p = 6.6e-11
R = 0.068, p = 0.15
R = −0.35, p = 5.5e-14
R = 0.12, p = 0.011
R = 0.0025, p = 0.96
R = 0.068, p = 0.15
R = −0.15, p = 0.0017
R = 0.28, p = 1.3e-09
R = −0.046, p = 0.33
R = −0.23, p = 8e-07
R = 0.56, p < 2.2e-16
R = 0.2, p = 3.1e-05
R = −0.02, p = 0.68
R = −0.046, p = 0.33
R = 0.21, p = 5.8e-06
R = −0.096, p = 0.043
R = −0.06, p = 0.21
R = 0.1, p = 0.03
R = −0.19, p = 5.4e-05

B

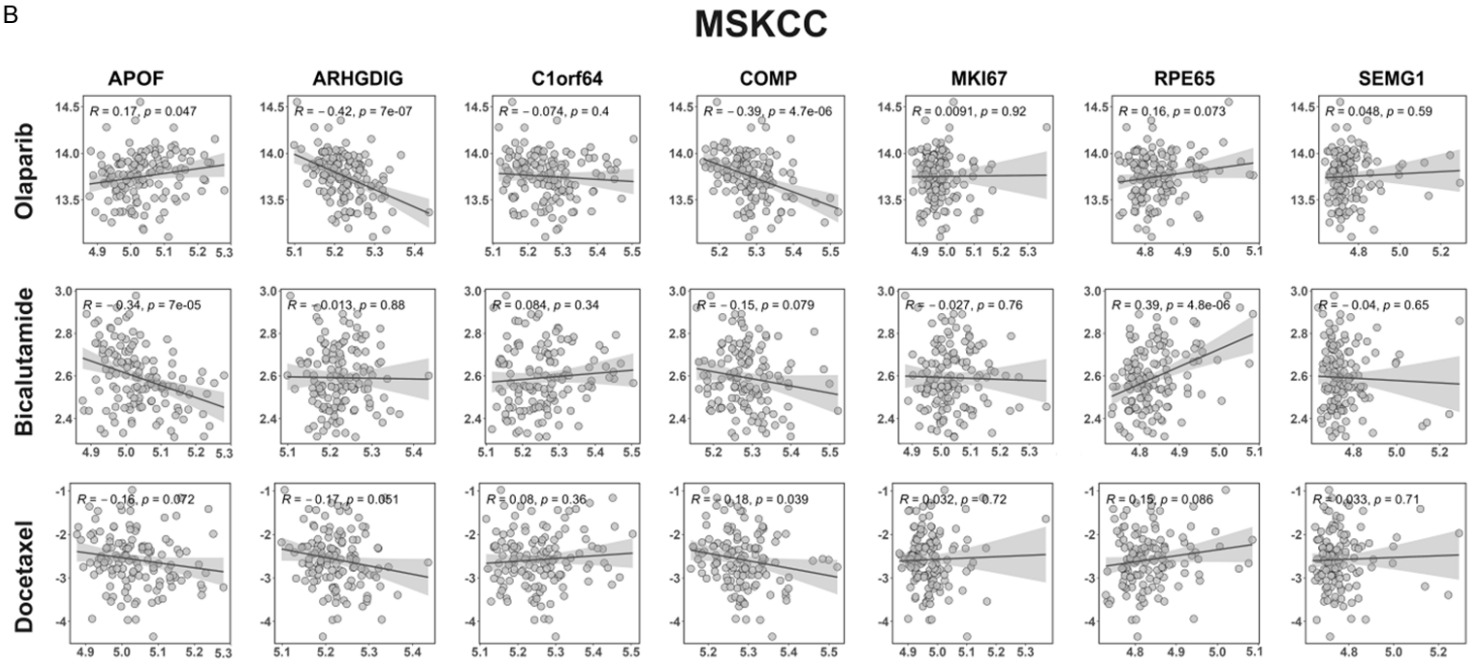

MSKCC
APOF
ARHGDIG
C1orf64
COMP
MKI67
RPE65
SEMG1
Olaparib
Bicalutamide
Docetaxel
R = 0.17, p = 0.047
R = −0.42, p = 7e-07
R = −0.074, p = 0.4
R = −0.39, p = 4.7e-06
R = 0.0091, p = 0.92
R = 0.16, p = 0.073
R = 0.048, p = 0.59
R = −0.34, p = 7e-05
R = −0.013, p = 0.88
R = 0.084, p = 0.34
R = −0.15, p = 0.079
R = −0.027, p = 0.76
R = 0.39, p = 4.8e-06
R = −0.04, p = 0.65
R = −0.16, p = 0.072
R = −0.17, p = 0.051
R = 0.08, p = 0.36
R = −0.18, p = 0.039
R = 0.032, p = 0.72
R = 0.15, p = 0.086
R = 0.033, p = 0.71

Figure S3. The correlation of gene expression and olaparib, docetaxel and bicalutamide responses in two cohorts. A. The correlations of APOF, ARHGDIG, Clorf64, COMP, MKI67, RPE65 and SEMG1 expressions and olaparib, docetaxel and bicalutamide responses in the TCGA cohort. B. The correlations of APOF, ARHGDIG, Clorf64, COMP, MKI67, RPE65 and SEMG1 expressions and olaparib, docetaxel and bicalutamide responses in the MSKCC cohort.

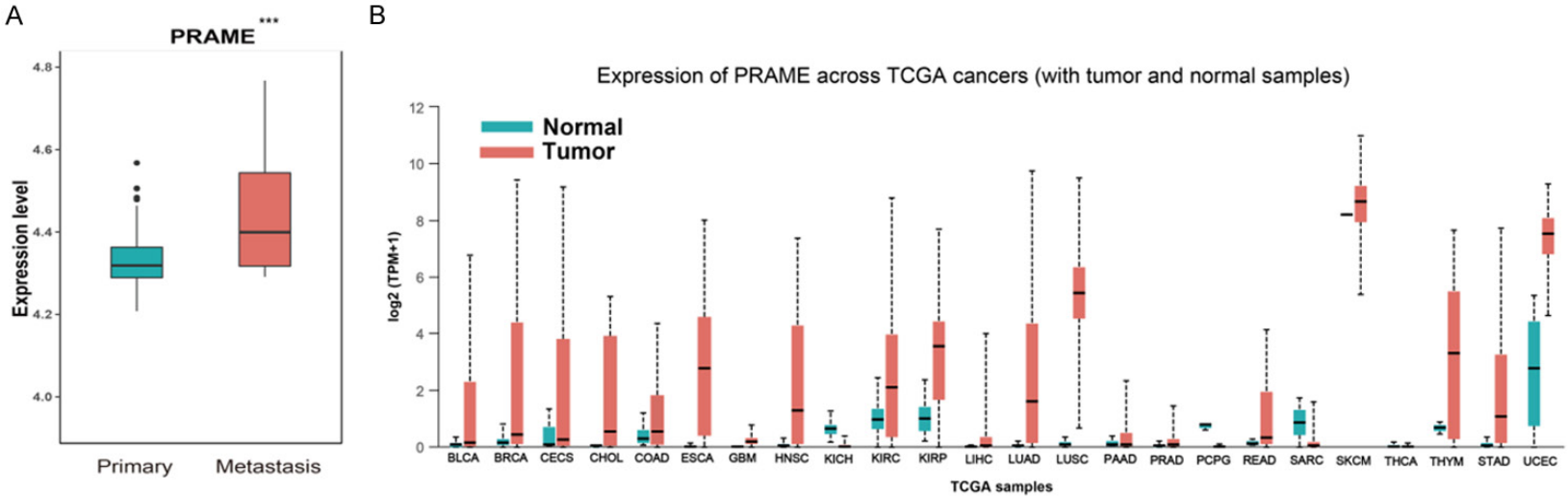


Figure S4. The expression of PRAME in metastasis and primary patients and in TCGA cancers. A. The PRAME expression in metastasis and primary patients in the MSKCC cohort. B. The PRAME expression in the TCGA cohort tumor and paired normal tissues.