



Establishment of Novel DNA Methylation-Based Prostate Cancer Subtypes and a Risk-Predicting Eight-Gene Signature

Enchong Zhang¹, Fujisawa Shiori², Oscar YongNan Mu³, Jieqian He¹, Yuntian Ge¹, Hongliang Wu⁴, Mo Zhang^{1*} and Yongsheng Song^{1*}

¹ Department of Urology, Shengjing Hospital of China Medical University, Shenyang, China, ² Department of Breast and Endocrine Surgery, Tohoku University Hospital, Sendai, Japan, ³ PANASIAUSMLE, Association of Asian Medical Graduates, Toronto, ON, Canada, ⁴ Department of Spine and Joint Surgery, Shengjing Hospital of China Medical University, Shenyang, China

OPEN ACCESS

Edited by:

Tao Huang,
Shanghai Institute of Nutrition
and Health (CAS), China

Reviewed by:

Heejuon Chae,
Sookmyung Women's University,
South Korea
Siyuan Han,
The First Affiliated Hospital of China
Medical University, China

*Correspondence:

Mo Zhang
peterzhang623@gmail.com
Yongsheng Song
sysuro@163.com;
songys@sj-hospital.org

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 05 January 2021

Accepted: 02 February 2021

Published: 23 February 2021

Citation:

Zhang E, Shiori F, Mu OY, He J,
Ge Y, Wu H, Zhang M and Song Y
(2021) Establishment of Novel DNA
Methylation-Based Prostate Cancer
Subtypes and a Risk-Predicting
Eight-Gene Signature.
Front. Cell Dev. Biol. 9:639615.
doi: 10.3389/fcell.2021.639615

Prostate cancer (PCa) is the most common malignant tumor affecting males worldwide. The substantial heterogeneity in PCa presents a major challenge with respect to molecular analyses, patient stratification, and treatment. Least absolute shrinkage and selection operator was used to select eight risk-CpG sites. Using an unsupervised clustering analysis, called consensus clustering, we found that patients with PCa could be divided into two subtypes (Methylation_H and Methylation_L) based on the DNA methylation status at these CpG sites. Differences in the epigenome, genome, transcriptome, disease status, immune cell composition, and function between the identified subtypes were explored using The Cancer Genome Atlas database. This analysis clearly revealed the risk characteristics of the Methylation_H subtype. Using a weighted correlation network analysis to select risk-related genes and least absolute shrinkage and selection operator, we constructed a prediction signature for prognosis based on the subtype classification. We further validated its effectiveness using four public datasets. The two novel PCa subtypes and risk predictive signature developed in this study may be effective indicators of prognosis.

Keywords: prostate cancer, DNA methylation, predictive signature, prognosis, *RND3*, hypermethylation, systems biology

INTRODUCTION

As the most common cancer in males, prostate cancer (PCa) is a major public health threat (Siegel et al., 2020). Based on the latest global cancer data from the World Health Organization (WHO)¹, the age-standardized rate for PCa ranks second, 5-year prevalence ranks first, and age-standardized mortality ranks sixth. Despite the ongoing development of therapeutic strategies, the heterogeneity of PCa contributes to treatment failure (Chang et al., 2014; Peng et al., 2018). Therefore, it is necessary to identify subtypes of PCa during diagnosis and treatment.

During the process of DNA methylation, methyl groups are added to CpG islands on the DNA molecule. Hypermethylation acts on promoters and could lead to gene silencing, whereas hypomethylation is associated with chromosomal instability and a loss of imprinting

¹<https://gco.iarc.fr/>

(Daura-Oller et al., 2009). In many diseases, including cancer, abnormal hypermethylation on gene promoters, which could be inherited by daughter cells, has been detected (Wang and Lei, 2018). Abnormal DNA methylation status is now considered a significant determinant of cancer development. Abundant stable DNA methylation in the genome is a candidate for diagnosis and treatment (Mikeska and Craig, 2014). Accordingly, the use of DNA methylation status to divide PCa into subtypes may provide important new insights.

Novel methylation-based subtypes have been reported in PCa. For example, the Cancer Genome Atlas Research Network (Abeshouse et al., 2015) conducted a comprehensive molecular analysis of 333 primary prostate carcinomas and identified different subtypes with highly diverse genomic, epigenomic, and transcriptomic patterns. In particular, an unsupervised hierarchical cluster analysis of the 5,000 most variable hypermethylated CpG sites revealed four epigenetically distinct subtypes of PCa. In another multicenter study, a new epigenetic CpG methylator phenotype in advanced PCa was reported, and this subtype is characterized by hypermethylation both within and outside CpG sites, shores, and shelves (Zhao et al., 2020). These important studies have improved our understanding of the DNA methylation landscape in PCa, providing a reference for future research. Recent studies have focused on revealing associations of novel DNA methylation subtypes with driver events in the genome and transcriptome. Utilizing a different approach, in this study, we identified DNA methylation subtypes based on clinical outcomes, with further analyses linking these subtypes with molecular mechanisms. In particular, we screened out eight CpG sites in which the DNA methylation status was associated with prognosis in PCa and identified two subtypes based on these DNA methylation statuses. Thereafter, differences in the epigenome, genome, transcriptome, disease status, immune cell infiltration, and function between these two subtypes were explored. Lastly, genes related to the high-risk subtype were selected and screened to construct an eight-gene signature with the ability to predict prognosis. The effectiveness of the signature was validated using four public datasets.

MATERIALS AND METHODS

Data Processing

RNA-seq data (in the form of HTSeq-Counts and HTSeq-FPKM), DNA methylation (450 K), somatic variation, copy number alterations (CNA), and clinical information for patients with PCa were downloaded from The Cancer Genome Atlas (TCGA) database² (Blum et al., 2018). The gene annotation file was downloaded from the Ensembl database³ (Howe et al., 2020). RNA-seq data in FPKM were converted to TPM. The TPM and β -values for CpG sites were quantile-normalized. In total, 477 patients with both the data described above (RNA-seq and methylation data) and complete clinical

information were included. Disease-related clinical information for these patients is provided in **Table 1**. To validate the effectiveness of the risk signature, gene expression profiles

TABLE 1 | The disease-related clinical information of patients with PCa included in the study.

Characteristics	Value
Patients (<i>n</i>)	477
Age (year), median (IQR)	62.0 (56.0–66.0)
PSA (ng/ml), median (IQR)	7.5 (5.1–11.4)
Pathological Gleason score, <i>n</i> (%)	
≤6	43 (9.0%)
7 (3+4)	143 (30.0%)
7 (4+3)	100 (21.0%)
8	56 (11.7%)
9~10	135 (28.3%)
Prior malignancy, <i>n</i> (%)	
No	450 (94.3%)
Yes	27 (5.7%)
Race, <i>n</i> (%)	
Asian	12 (2.5%)
Whit, American Indian or Alaska native	398 (83.4%)
Black or African American	55 (11.6%)
NA	12 (2.5%)
Residual tumor, <i>n</i> (%)	
R0	301 (63.1%)
R1	15 (3.1%)
R2	142 (29.8%)
Rx	5 (1.0%)
NA	14 (3.0%)
Clinical M, <i>n</i> (%)	
M0	437 (91.6%)
M1a or M1c	2 (0.4%)
NA	38 (8.0%)
Pathological T, <i>n</i> (%)	
T1c	2 (0.4%)
T2a	13 (2.7%)
T2b	10 (2.1%)
T2c	160 (33.5%)
T3a	151 (31.7%)
T3b	129 (27.0%)
T4	9 (1.9%)
NA	3 (0.7%)
Pathological N, <i>n</i> (%)	
N0	329 (69.0%)
N1	78 (16.4%)
NA	70 (14.6%)
Diagnostic CT or MRI, <i>n</i> (%)	
No evidence of extraprostatic extension	196 (41.1%)
Equivocal	6 (1.3%)
Extraprostatic extension localized	22 (4.6%)
Extraprostatic extension	9 (1.9%)
NA	244 (51.1%)
Outcome, <i>n</i> (%)	
DFS	53 (11.1%)
Disease free	424 (88.9%)

PCa, prostate cancer; PSA, prostate-specific antigen; DFS, disease-free survival; IQR, interquartile range; NA, not analyzed.

²<https://portal.gdc.cancer.gov/>

³<http://asia.ensembl.org/index.html>

and clinical data from four public datasets were used, i.e., GSE70769 and GSE116918 from the Gene Expression Omnibus (GEO) database⁴ (Barrett et al., 2013; Ross-Adams et al., 2015; Jain et al., 2018) and DKF2018 and MSKCC2010 from cBioPortal for Cancer Genomics⁵ (Cerami et al., 2012; Gao et al., 2013). Information about these four datasets is provided in Table 2.

Identification of DNA Methylation-Based Subtypes

We previously identified 120 CpG sites that were differentially methylated between PCa and normal prostate tissues and were significantly associated with disease-free survival (DFS) (Zhang et al., 2020b). Least absolute shrinkage and selection operator (LASSO) regression enables variable selection and regularization, while fitting the generalized linear model. Therefore, LASSO regression was used to reduce the number of CpG sites as the input for subtype identification using the glmnet R package (Engelbrechtsen and Bohlin, 2019). After the CpG sites were screened, the β values for these sites in 477 patients were used as inputs for consensus clustering, an unsupervised clustering analysis. Consensus clustering was performed using the ConsensusClusterPlus R package (Monti et al., 2003) and the following operating parameters: maxK = 10, reps = 1000, pItem = 0.8, pFeature = 1, clusterAlg = hc, and distance = pearson. A heatmap was generated to visualize the methylation status of the subtypes using the pheatmap R package (Kolde and Kolde, 2015). Furthermore, a survival analysis of the subtypes was performed by the log rank test using the survival R package (Therneau and Lumley, 2014). To further demonstrate the differences between subtypes, a principal component analysis was performed using the 477 patients. Finally, correlations between disease-related clinical information and subtype status were evaluated by the Mann–Whitney *U* test, χ^2 test, or Fisher's exact test.

Single Nucleotide Variation Between Subtypes

Simple nucleotide variants were compared between subtypes using the GenVisR R package (Skidmore et al., 2016). Genes

⁴<https://www.ncbi.nlm.nih.gov/geo/>

⁵<https://www.cbioportal.org>

TABLE 2 | Information of the four publicly available independent validation datasets.

Dataset	Sample size	Transcriptome platform	Tissue
DKFZ2018	82	Illumina HiSeq 2000 (RNAseq)	Fresh frozen
GSE70769	90	Illumina humanHT-12 V4.0	Fresh frozen
GSE1166918	248	ADXPCv1a520642 Affymetrix human	Formalin-fixed Paraffin-embedded
MSKCC2010	140	Affymetrix human Exon 1.0 ST array	Fresh frozen

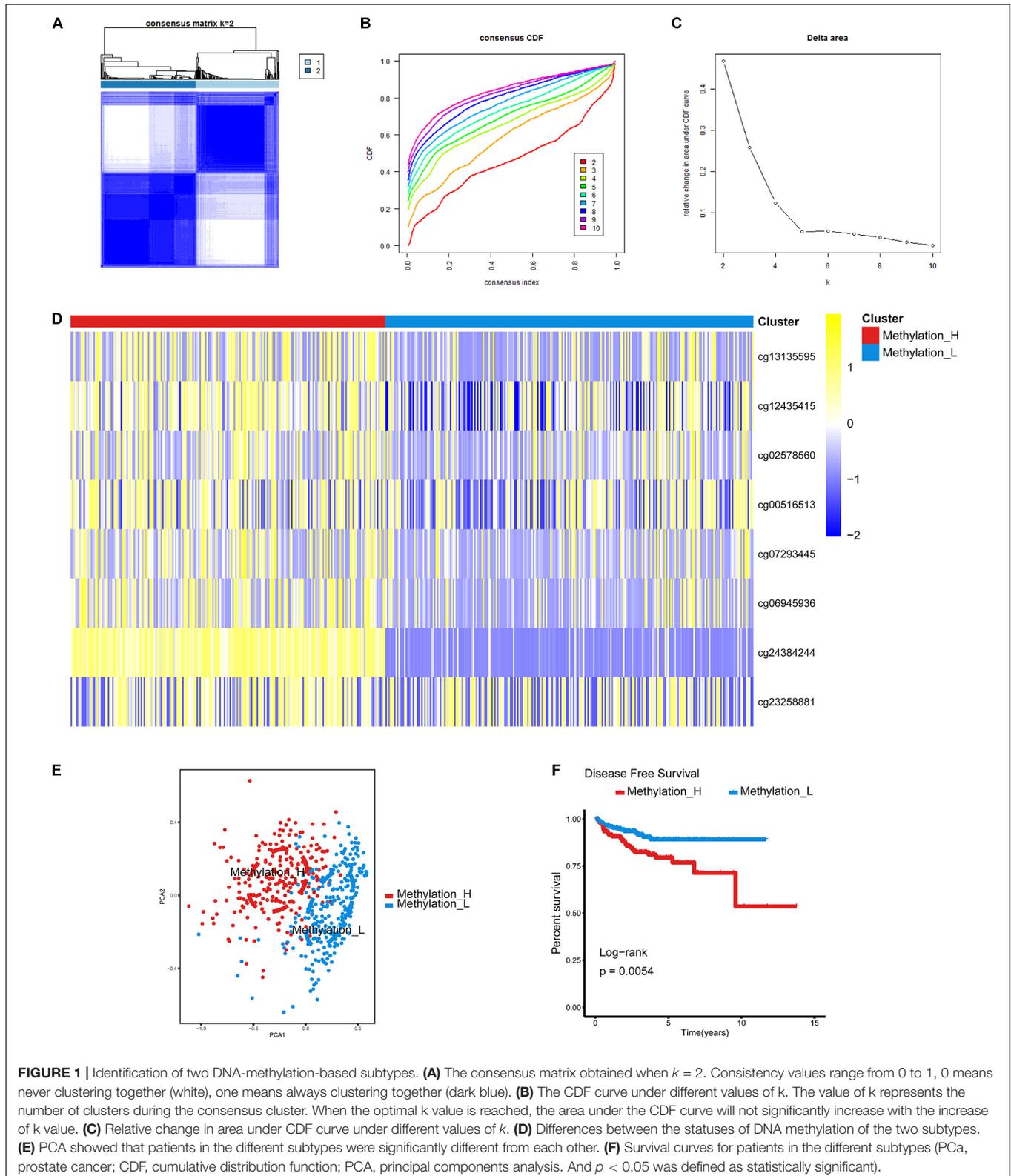
TABLE 3 | The association between subtypes and disease-related clinical information of PCa.

Clinicopathologic variables	Consensus clusters		P
	Methylation_H (n = 220)	Methylation_L (n = 257)	
Age (year), median (IQR)	63.0 (57.8–67.0)	60.0 (55.0–65.0)	0.001a
PSA (ng/ml), median (IQR)	8.0 (5.8–12.9)	6.8 (4.6–10.1)	< 0.001a
Pathological Gleason score, n (%)			< 0.001b
≤6	11 (5.0%)	32 (12.5%)	
7 (3+4)	53 (24.1%)	90 (35.0%)	
7 (4+3)	45 (20.5%)	55 (21.4%)	
8	26 (11.8%)	30 (11.7%)	
9~10	85 (38.6%)	50 (19.4%)	
Prior malignancy, n (%)			0.828b
No	207 (94.1%)	243 (94.6%)	
Yes	13 (5.9%)	14 (5.4%)	
Race, n (%)			0.529b
Asian	7 (3.2%)	5 (2.0%)	
Whit, American Indian or Alaska native	181 (82.3%)	217 (84.4%)	
Black or African American	28 (12.7%)	27 (10.5%)	
NA	4 (1.8%)	8 (3.1%)	
Residual tumor, n (%)			0.005b
R0	126 (57.3%)	175 (68.1%)	
Rx/R1/R2	90 (40.9%)	72 (28.0%)	
NA	4 (3.0%)	10 (3.9%)	
Clinical M, n (%)			0.226c
M0	207 (94.1%)	230 (89.5%)	
M1a or M1c	2 (0.9%)	0 (0.0%)	
NA	11 (5.0%)	27 (10.5%)	
Pathological T, n (%)			0.013c
T1c	0 (0.0%)	2 (0.8%)	
T2a	3 (1.4%)	10 (3.9%)	
T2b	3 (1.4%)	7 (2.7%)	
T2c	63 (28.6%)	97 (37.7%)	
T3a	71 (32.3%)	80 (31.1%)	
T3b	73 (33.2%)	56 (21.8%)	
T4	6 (2.7%)	3 (1.2%)	
NA	1 (0.4%)	2 (0.8%)	
Pathological N, n (%)			0.289b
N0	151 (68.6%)	178 (69.3%)	
N1	41 (18.6%)	37 (14.4%)	
NA	28 (12.8%)	42 (16.3%)	
Diagnostic CT or MRI, n (%)			0.135c
No evidence of extraprostatic extension	95 (43.2%)	101 (39.3%)	
Equivocal	5 (2.3%)	1 (0.4%)	
Extraprostatic extension localized	10 (4.5%)	12 (4.7%)	
Extraprostatic extension	7 (3.2%)	2 (0.8%)	
NA	103 (46.8%)	141 (54.8%)	
Outcome, n (%)			0.002b
DFS	35 (8.3%)	18 (7.0%)	
Disease free	185 (91.7%)	239 (93.0%)	

P values were calculated by the Mann–Whitney *U* test (a), Chi-square test (b), or Fisher's exact test (c). PCa, prostate cancer; PSA, prostate-specific antigen; DFS, disease-free survival; IQR, interquartile range; NA, not analyzed.

with the top 10 mutation frequencies were displayed in a waterfall plot. According to the results of the waterfall plot, the difference between the mRNA levels of mutant and wild-type

Speckle-type POZ Protein (*SPOP*) was evaluated. Then, the mRNA levels of *SPOP* in different subtypes were compared by Wilcoxon's test.



Differences in Copy Number Alterations, TMPRSS2-ERG Fusion, and Androgen Receptor Scores Between Subtypes

To explore the difference in CNAs between subtypes, genes with significant differences in copy number between subtypes were identified by chi-squared tests. Among these genes, *RND3* was differentially expressed between the subtypes, as determined by a Wilcoxon test. The type and frequency of CNAs in *RND3* were explored. Furthermore, the relationship between CNA types and mRNA expression levels of *RND3* were evaluated by the Wilcoxon test.

Data from The Tumor Fusion Gene Data Portal database (<https://www.tumorfusions.org/>) were used to analyze the difference in *TMPRSS2-ERG* fusion gene expression between the subtypes (Yoshihara et al., 2015). Finally, the androgen receptor (AR) score in each subtype was compared by the Wilcoxon test. AR scores were obtained from the cBioPortal database (Cancer Genome Atlas Research Network, 2015).

Immune Cells in the Tumor Microenvironment in Each Subtype

RNA-seq data in TPM format were uploaded to CIBERSORTx⁶ (Newman et al., 2019) to evaluate the infiltration of 22 types of immune cells in the tumor microenvironment. The abundances

of these immune cells were compared between subtypes using a violin plot and the Wilcoxon test. Furthermore, survival curves were generated for these cells. The p-values for the survival analysis were calculated by a Cox regression and log-rank test using the survival R package (Therneau and Lumley, 2014).

Functional Enrichment Analysis of Subtypes

Fold change values for gene expression differences between the subtypes were used as the ranks in a gene set enrichment analysis (GSEA). To obtain fold changes, HTSeq-Counts were analyzed using the DESeq2 R package (Love et al., 2014). The hallmark gene set downloaded from the Molecular Signatures Database9 v7.1⁷ was used as the reference gene list in the GSEA (Subramanian et al., 2005; Liberzon et al., 2011). Finally, the GSEA was completed using the clusterProfiler R package (Yu et al., 2012).

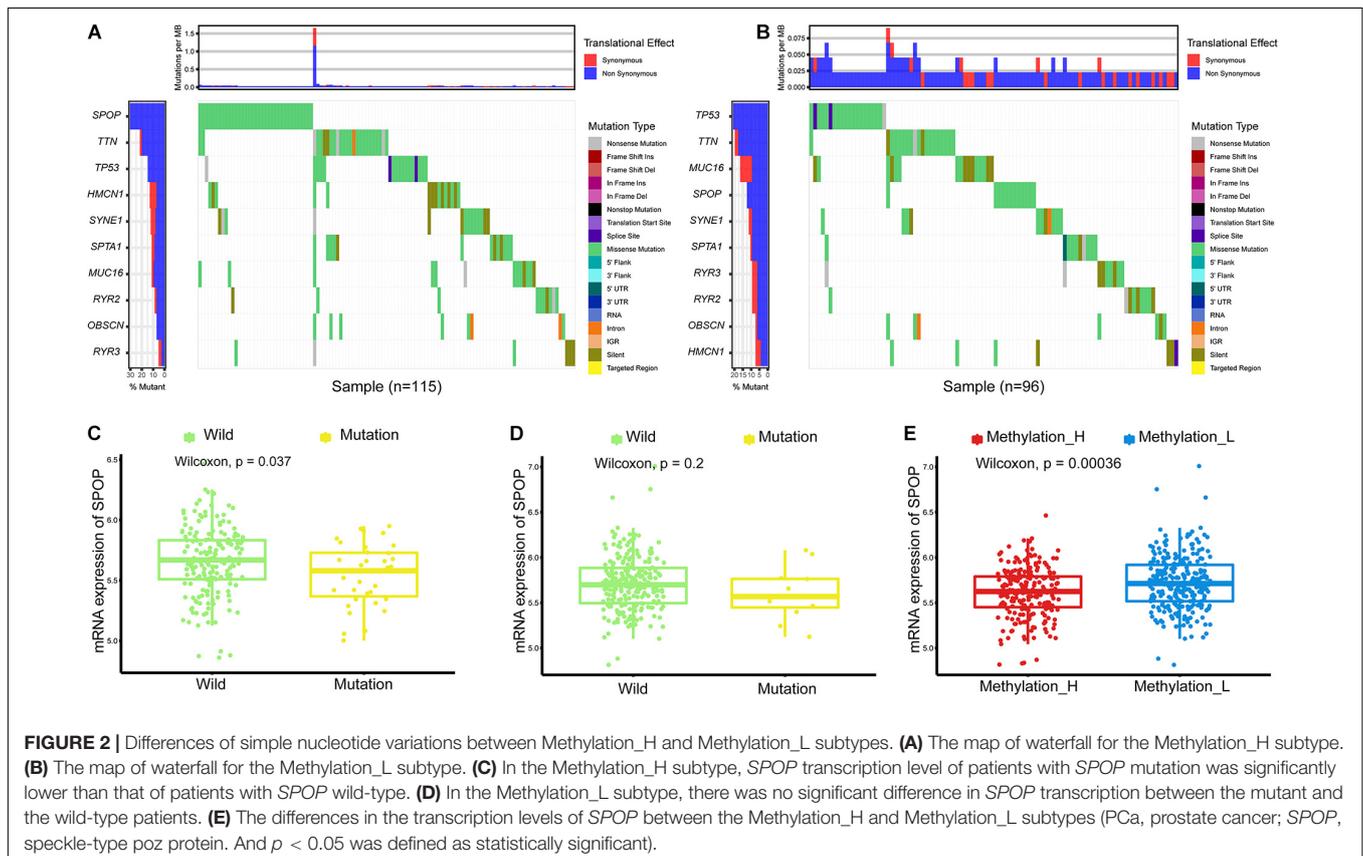
Furthermore, expression levels of genes that were crucial for PCA were compared between the subtypes by Wilcoxon tests.

Weighted Correlation Network Analysis of Subtypes

A weighted correlation network analysis (WGCNA) could be used find phenotype-associated gene modules

⁶<https://cibersortx.stanford.edu/>

⁷<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>



(Langfelder and Horvath, 2008; Li et al., 2019). Therefore, TPM values from RNA-seq data were used as the input for a WGCNA. Eight was the soft power threshold to construct a network that simultaneously satisfied a scale-free topology and high connectivity. Pearson correlation coefficients for the relationships between phenotypes and gene modules were determined. The phenotypes included PSA (prostate-specific antigen), Gleason score, and the subtypes. The gene module most closely associated with the high-risk phenotypes was identified. Differentially expressed genes (DEGs) between PCa and normal prostate tissues in the selected gene module were identified. The conditions for DEGs were logarithmic fold changes ($|\text{LFCs}| > 1$ and $p < 0.05$). Then, survival-associated genes were screened from the DEGs by Cox regression and log-rank tests. Finally, genes for LASSO were filtered out.

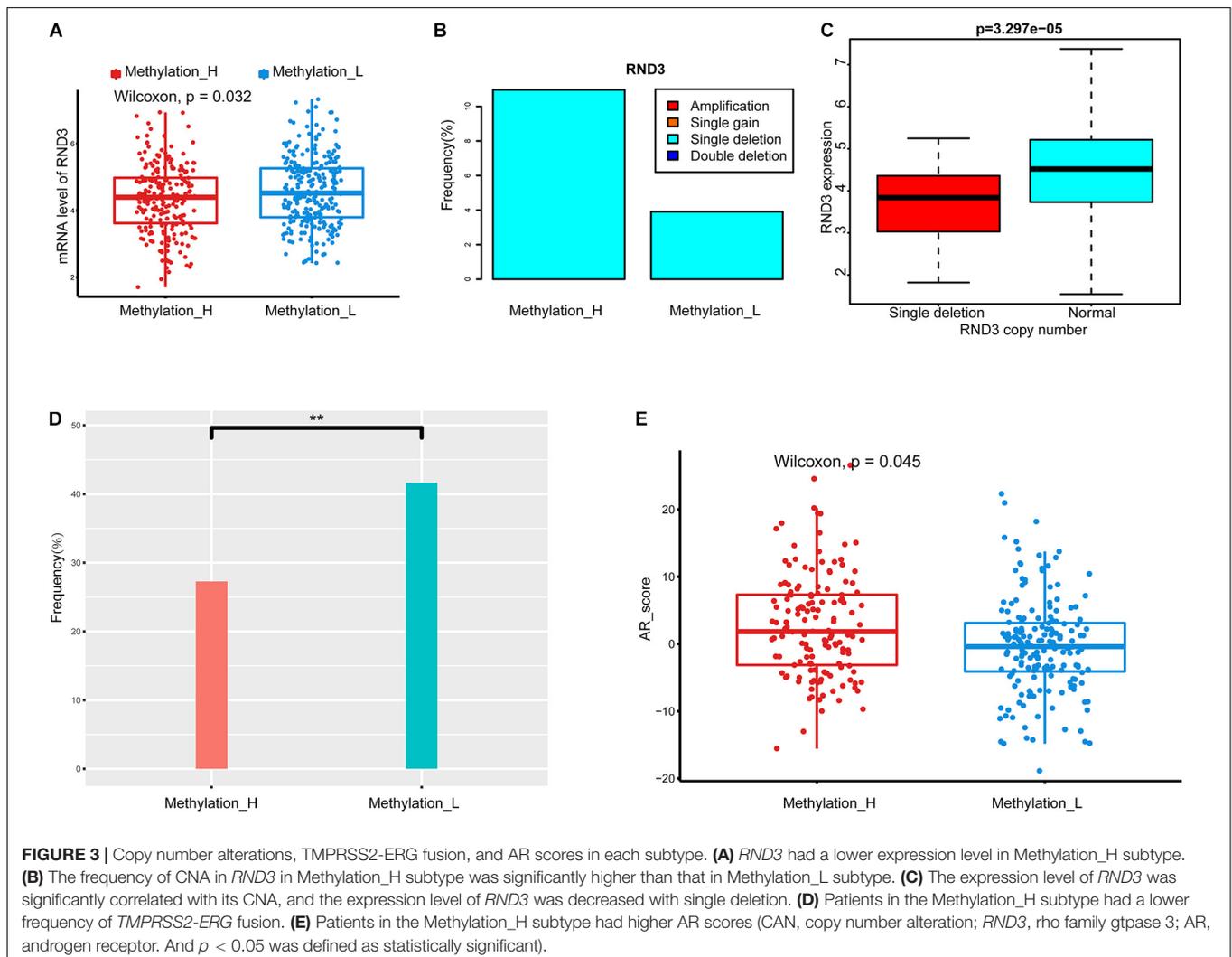
Identification of a Risk Signature in the Training Set

Before training, 477 patients were randomly divided into a training set and internal validation set using the caret R package

(Kuhn, 2008). Information for patients in the training set is provided in **Supplementary Table 1** and information for the internal validation set is provided in **Supplementary Table 2**. LASSO regression was used to construct a single signature for predicting prognosis with high performance (Svane et al., 2018). LASSO regression was applied to the training set; during the selection of genes, the C-index after 10-fold cross-validation reflected the effect of different screening strategies. Genes with the maximal C-index values were selected for the prognostic signature using the glmnet R package (Engelbrechtsen and Bohlin, 2019) with the following parameter settings: family = Cox, type.measure = C, parallel = TRUE, with default settings for other parameters. Furthermore, the difference in the risk score between subtypes identified and the relationship between the risk score and survival were evaluated. Comparisons were performed using the Wilcoxon test.

Predictive Accuracy of the Signature

First, time-dependent receiver operating characteristic (tdROC) curves were used to evaluate the predictive accuracy of the



signature in the training set, internal validation set, and external validation sets (DKFZ2018, GSE70769, GSE116918, and MSKCC2010) using the timeROC R package (Blanche and Blanche, 2013). Then, a survival analysis by Cox regression and the log-rank test was performed using these datasets. Survival curves were plotted using the Kaplan–Meier method and the survminer R package (Kassambara et al., 2017).

Univariate and multivariate Cox regression analyses were used to explore whether the risk score is an independent predictor of prognosis. Finally, the clinical diagnostic value of the signature was compared with that of clinical features (Gleason score and PSA) by a decision curve analysis (DCA) (Van Calster et al., 2018). DCA is used to compare prediction models that incorporate clinical outcomes; it requires only the dataset on which the models are tested and can be applied to models with either continuous or dichotomous results (Zhang et al., 2020a).

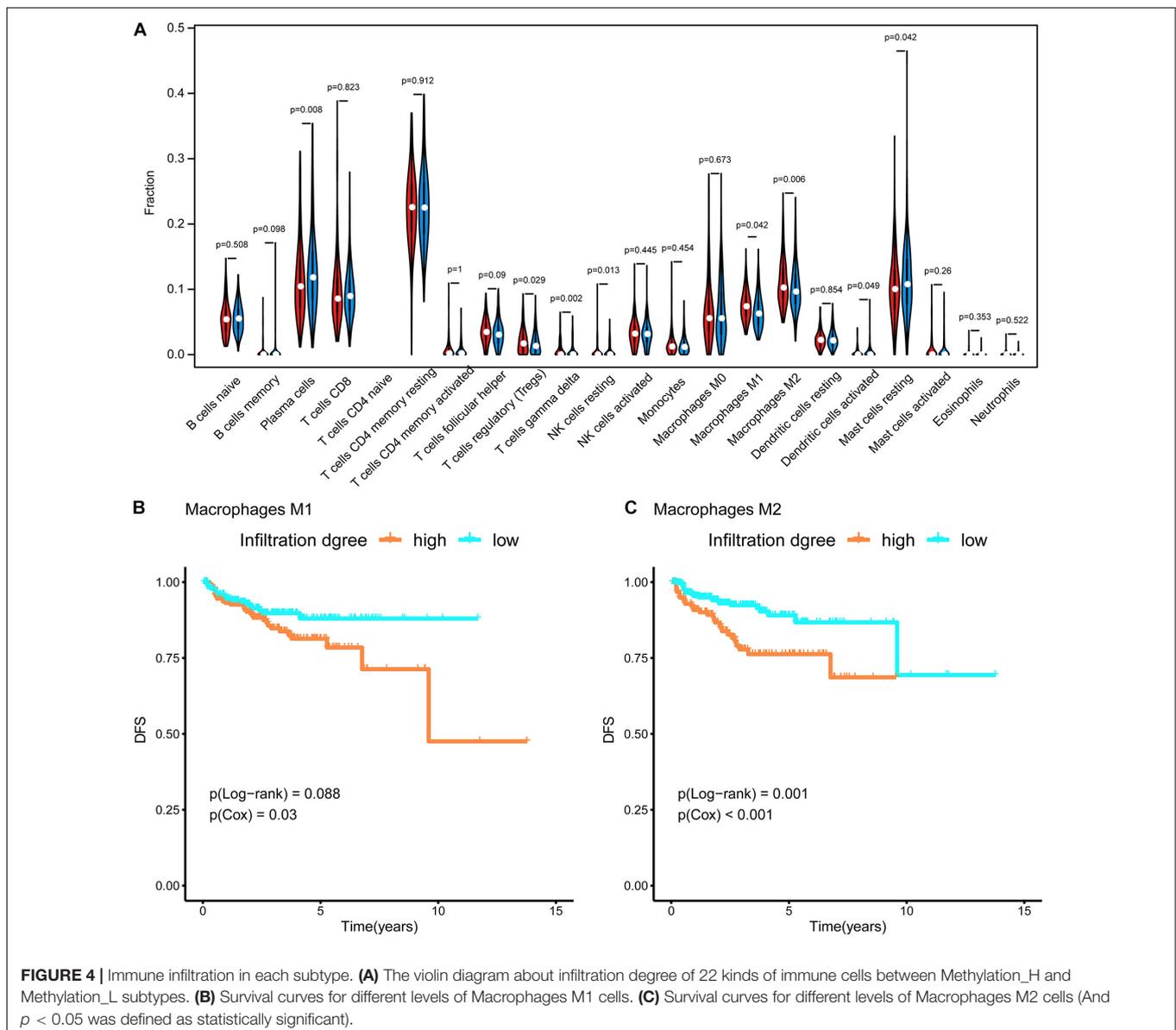
Statistical Analysis

R 3.6.3 was used for all statistical analyses. Values of $p < 0.05$ were defined as statistically significant. In the survival analysis, the survival outcome was defined as DFS or biochemical recurrence-free survival (BCR) based on clinical records.

RESULTS

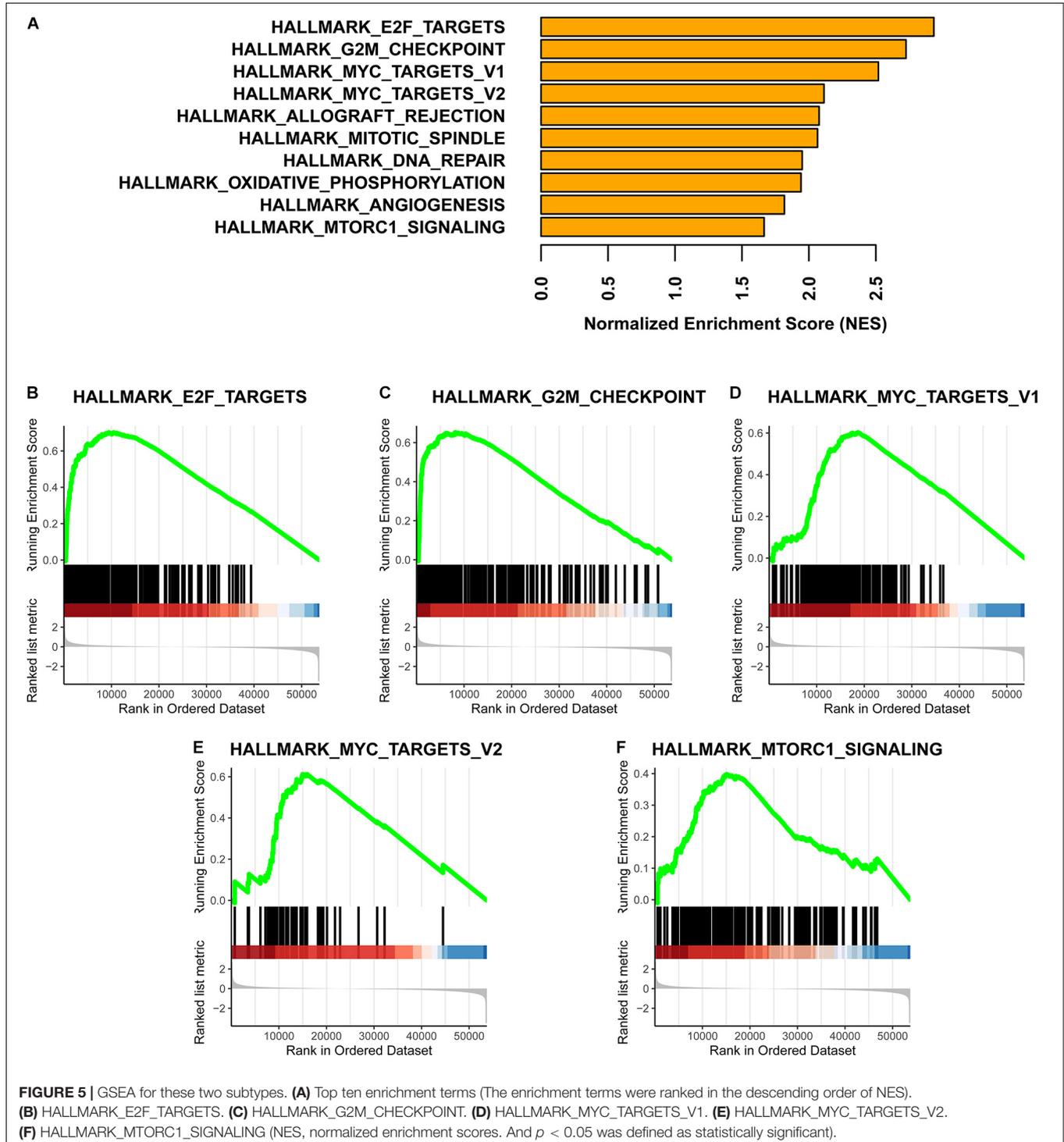
Identification of Two DNA Methylation-Based Subtypes

The cumulative distribution function (CDF) and relative change in the area under the CDF curve are shown in **Figures 1A,B**, respectively. According to Monti et al. (2003), the optimal k value is determined by a number of factors. One of the criteria



is that when the optimal k value is reached, the area under the CDF curve will not increase significantly with increases in k . Therefore we first assumed that the optimal value in this study was set to $k = 5$, indicating that the cohort could be divided into up to five subtypes. However, one cluster consisted of only a single patient when $k = 4$ or 5 . Additionally, the cluster-consensus value for each cluster was not large enough under $k = 4$ or 5 (Supplementary Figure 1). Therefore, we focused on $k = 2$ or

3. For $k = 3$, patients in C1 showed a worse prognosis, and patients in C2 and C3 did not show an obvious difference in prognosis (Figure 1C). For $k = 2$, patients in Methylation_H had worse a prognosis than that of patients in Methylation_L (Figure 1D). Furthermore, the Methylation_H subtype and C1 subtype included the same patients and the Methylation_L subtype consisted of patients in the C2 and C3 subtypes. Groups for different k values are shown in Supplementary Table 3.



Ultimately, we identified two subtypes with a difference in prognosis. The consensus matrix is displayed in the form of a heatmap in **Figure 1E** and the consensus clustering analysis is summarized in **Supplementary Figure 1**. A principal component analysis revealed clear separation between the Methylation_H subtype and the Methylation_L subtype (**Figure 1F**). As shown in **Figure 1G**, these two subtypes had different levels of DNA methylation at CpG sites. We defined the hypermethylated subtype as Methylation_H and the hypo methylated subtype as Methylation_L. As shown in **Table 3**, age, PSA, Gleason score, residual tumor status, pathological T, and clinical outcome of patients between Methylation_H and Methylation_L subtypes are different significantly.

Single Nucleotide Variations in Methylation_H and Methylation_L

Single nucleotide variations in genes with the top 10 mutation frequencies in these two subtypes are shown in **Figures 2A,B**. We found that the frequency of single nucleotide variations in *SPOP* was higher in the Methylation_H subtype than in the Methylation_L subtype. *SPOP* is one of the most frequently mutated genes in primary PCa. Based on the tumor-suppressive role of *SPOP* in PCa and the results of loss-of-function assays, *SPOP* mutations are expected to include the invasion and proliferation of PCa cells (Barbieri et al., 2012; An et al., 2014). Furthermore, within the Methylation_H subtype, mRNA expression levels of *SPOP* in patients with mutations were significantly lower than those in patients with wild-type *SPOP* (**Figure 2C**). However, this pattern was not observed in the Methylation_L subtype (**Figure 2D**). Finally, we found that the mRNA expression level of *SPOP* in the Methylation_H subtype was significantly lower than that in the Methylation_L subtype (**Figure 2E**). These results supported the risk characteristics of Methylation_H.

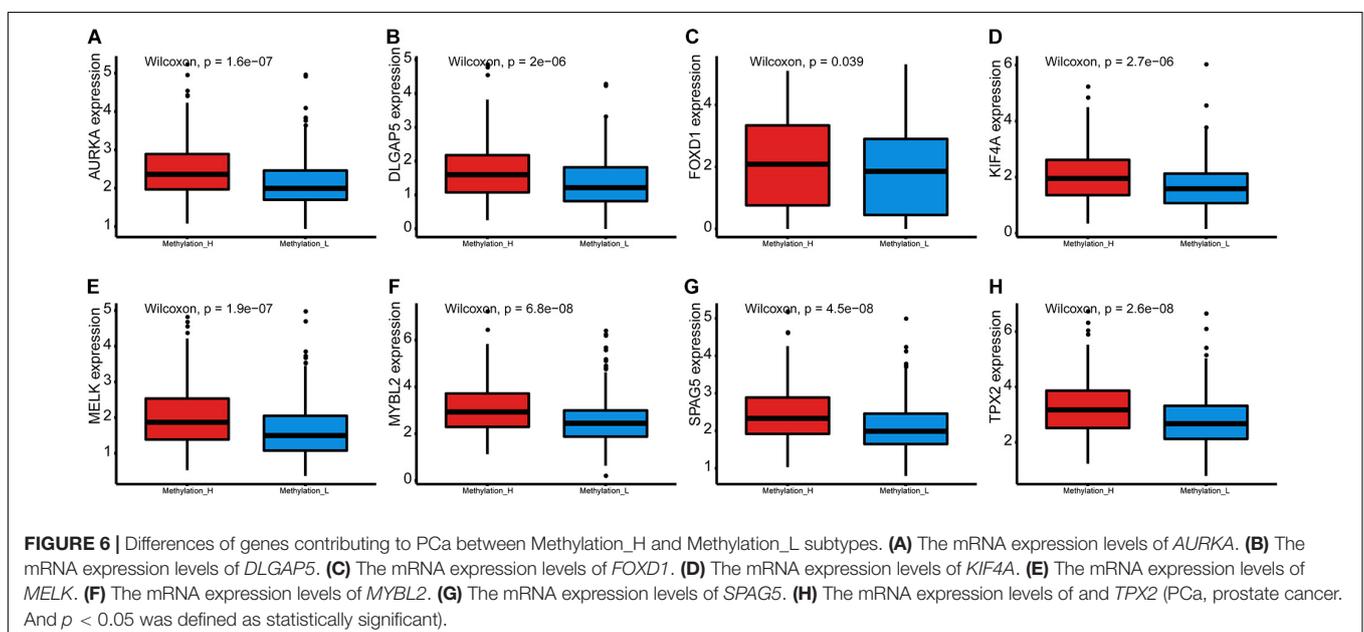
Copy Number Alterations of *RND3*, *TMPRSS2-ERG* Fusion, and Androgen Receptor Scores Between the Methylation_H and Methylation_L Subtypes

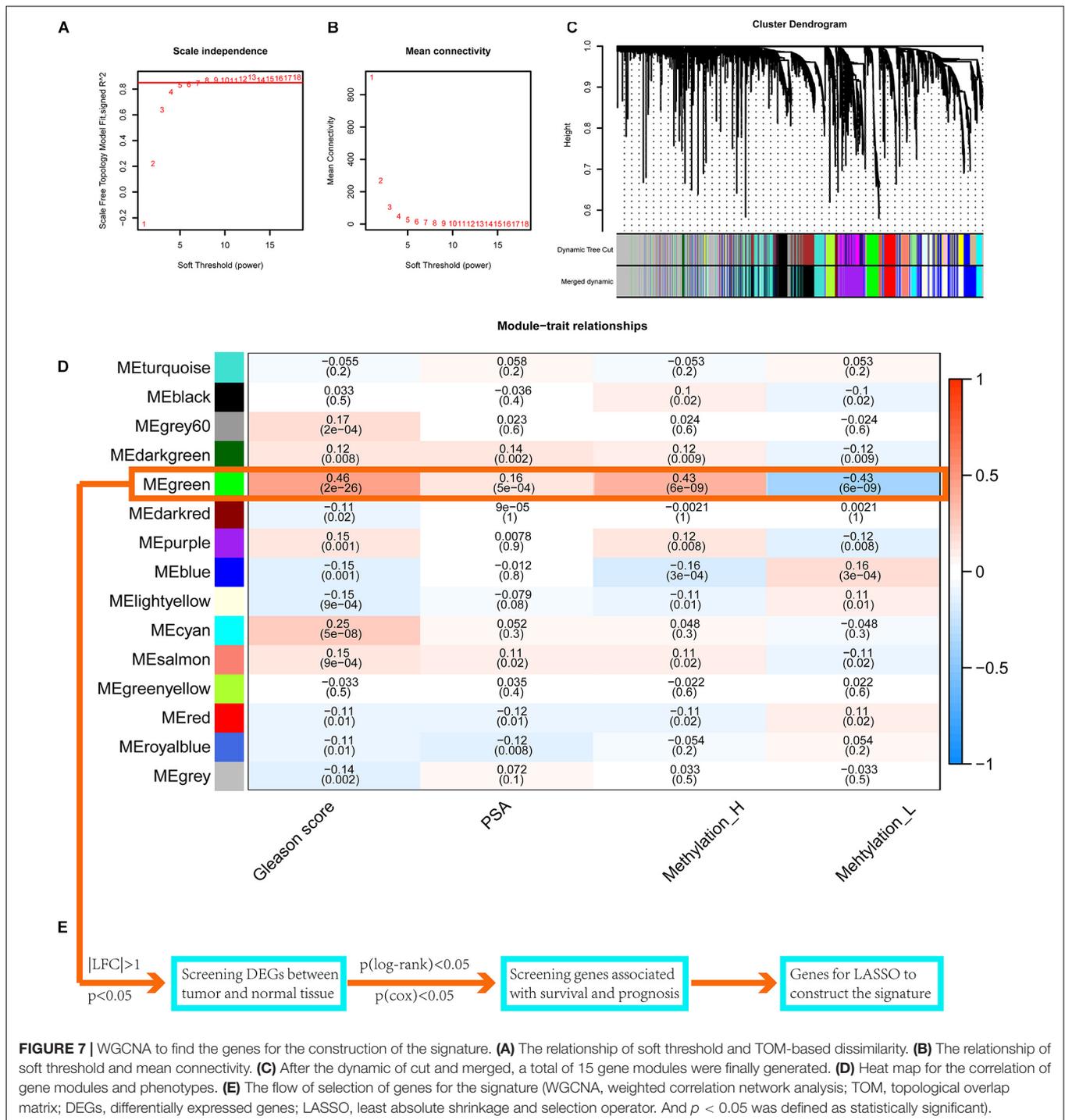
We identified a group of genes with a significant difference in copy number between subtypes. Among these genes, *RND3*, also called *RhoE*, is a tumor suppressor that is downregulated early in the development of PCa (Bektic et al., 2005). Interestingly, the expression of *RND3* was significantly lower in the Methylation_H subtype than that in the Methylation_L subtype (**Figure 3A**). For *RND3*, the type of CNA and the frequency of CNAs are summarized in **Figure 3B**. In 477 patients, a single-copy deletion was the only CNA detected in *RND3*. This deletion was more frequent in the Methylation_H subtype. Additionally, *RND3* was significantly down-regulated in cases with the single-copy deletion (**Figure 3C**).

The frequency of the *TMPRSS2-ERG* fusion was significantly lower in the Methylation_H subtype than in the Methylation_L subtype (**Figure 3D**). Some studies have revealed that the *TMPRSS2-ERG* fusion is related to the invasiveness of PCa and a higher Gleason score (Perner et al., 2006; Mehra et al., 2007; Rajput et al., 2007; Cheville et al., 2008). However, other studies have reported that the *TMPRSS2-ERG* fusion is not related to prognosis in PCa (Yoshimoto et al., 2006; Tu et al., 2007; Darnel et al., 2009). Furthermore, AR scores for patients in the Methylation_H subtype were higher than those of patients in the Methylation_L subtype (**Figure 3E**).

Immune Cells in the Tumor Microenvironment in Each Subtype

In **Figure 4A**, the difference in the immune cell composition in the tumor microenvironment of each subtype is displayed





in the form of a violin plot. Plasma cells and resting mast cells were significantly less abundant in the Methylation_H subtype and regulatory T cells (Tregs), M1 macrophages, and M2 macrophages were significantly more abundant in the Methylation_H subtype than in the Methylation_L subtype. Among these immune cells, greater M1 and M2 macrophage infiltration in the tumor microenvironment was related to a worse prognosis in PCA (Figures 4B,C).

Functional Enrichment Analysis of Each Subtype

Based on a GSEA, we ranked enriched terms in descending order of normalized enrichment scores. The top ten enriched terms are displayed in the Figure 5A. Among these terms, HALLMARK_E2F_TARGETS, HALLMARK_G2M_CHECKPOINT, HALLMARK_MYC_TARGETS_V1, HALLMARK_MYC_TARGETS_V2, and HALLMARK_

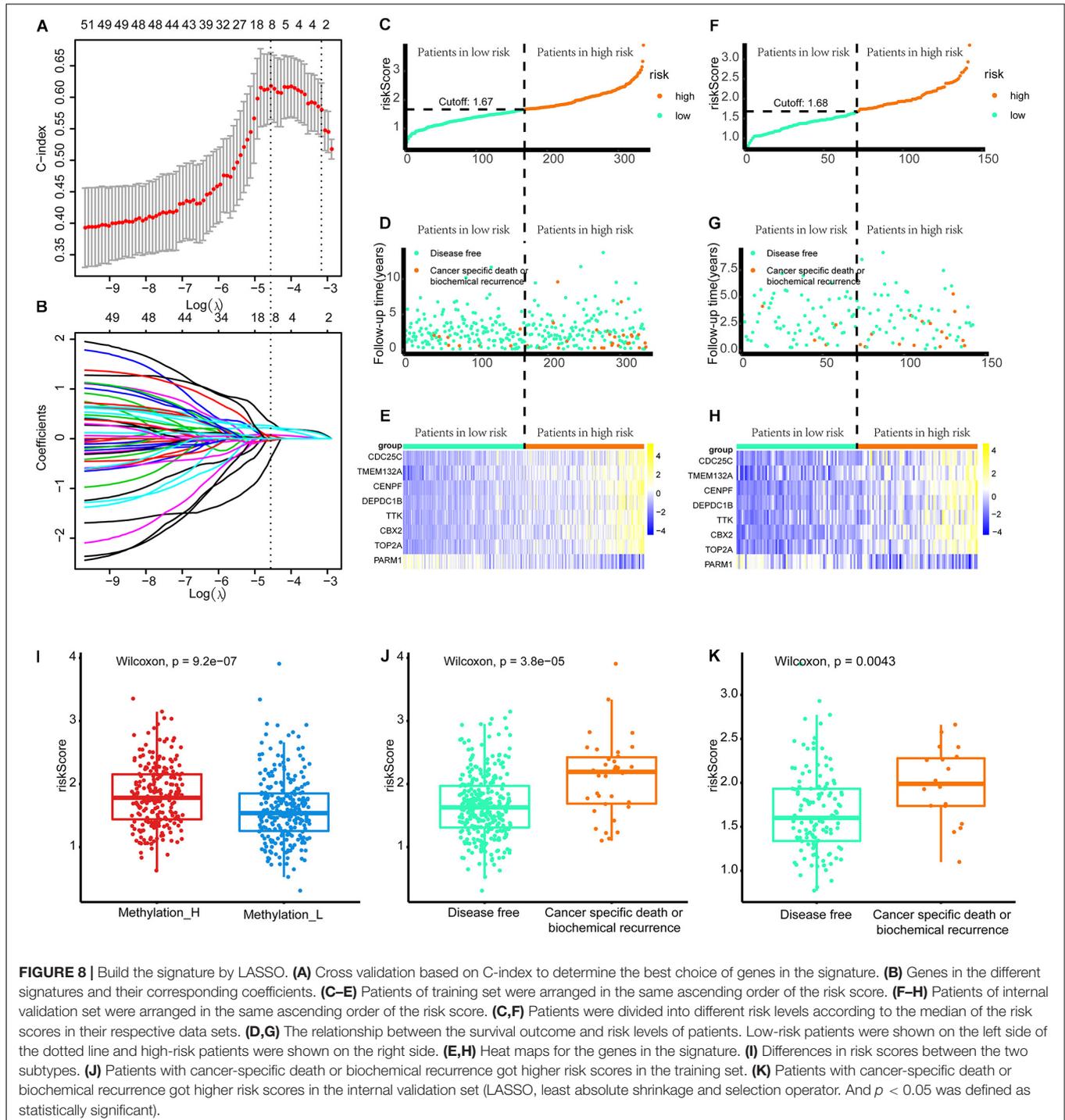
MTORC1_SIGNALING were enriched in the Methylation_H subtype (**Figures 5B–F**).

Furthermore, the mRNA expression levels of *AURKA*, *DLGAP5*, *FOXD1*, *KIF4A*, *MELK*, *MYBL2*, *SPAG5*, and *TPX2* were significantly higher in the Methylation_H subtype than in the Methylation_L subtype (**Figure 6**). These genes have all been reported to facilitate the development and progression of PCa (Kuner et al., 2013; Zhang et al., 2016, 2020b;

Akamatsu et al., 2018; Hewit et al., 2018; Zou et al., 2018; Cao et al., 2020; Li et al., 2020).

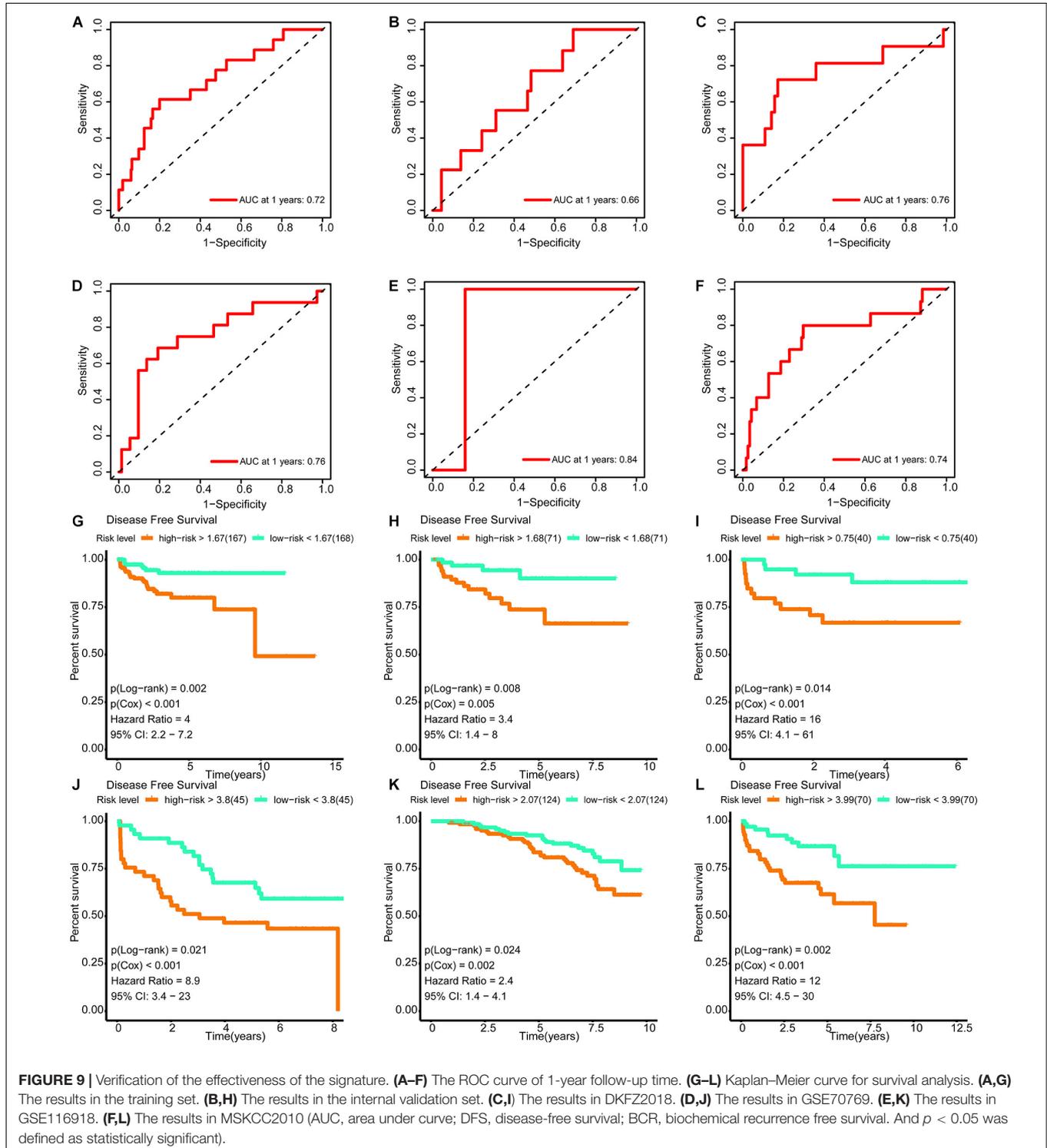
WGCNA for the Identification of a Key Gene Module

Setting eight as the soft threshold, the independence of the scale-free topology and the mean connectivity in each



module were sufficient (Figures 7A,B). After the dynamic cut and merge process, 15 gene modules were generated (Figure 7C). Among these, the MEgreen module was positively correlated with the PSA level, Gleason score, and the Methylation_H subtype (Figure 7D) and was negatively correlated with the Methylation_L subtype. Therefore,

genes in this module were related to the development and progression of PCa. According to the flow diagram in Figure 7E, we then screened out the DEGs between PCa and normal prostate tissues in the MEgreen module. Survival-associated DEGs were further identified for LASSO regression.



Construction of the Gene-Based Risk Signature

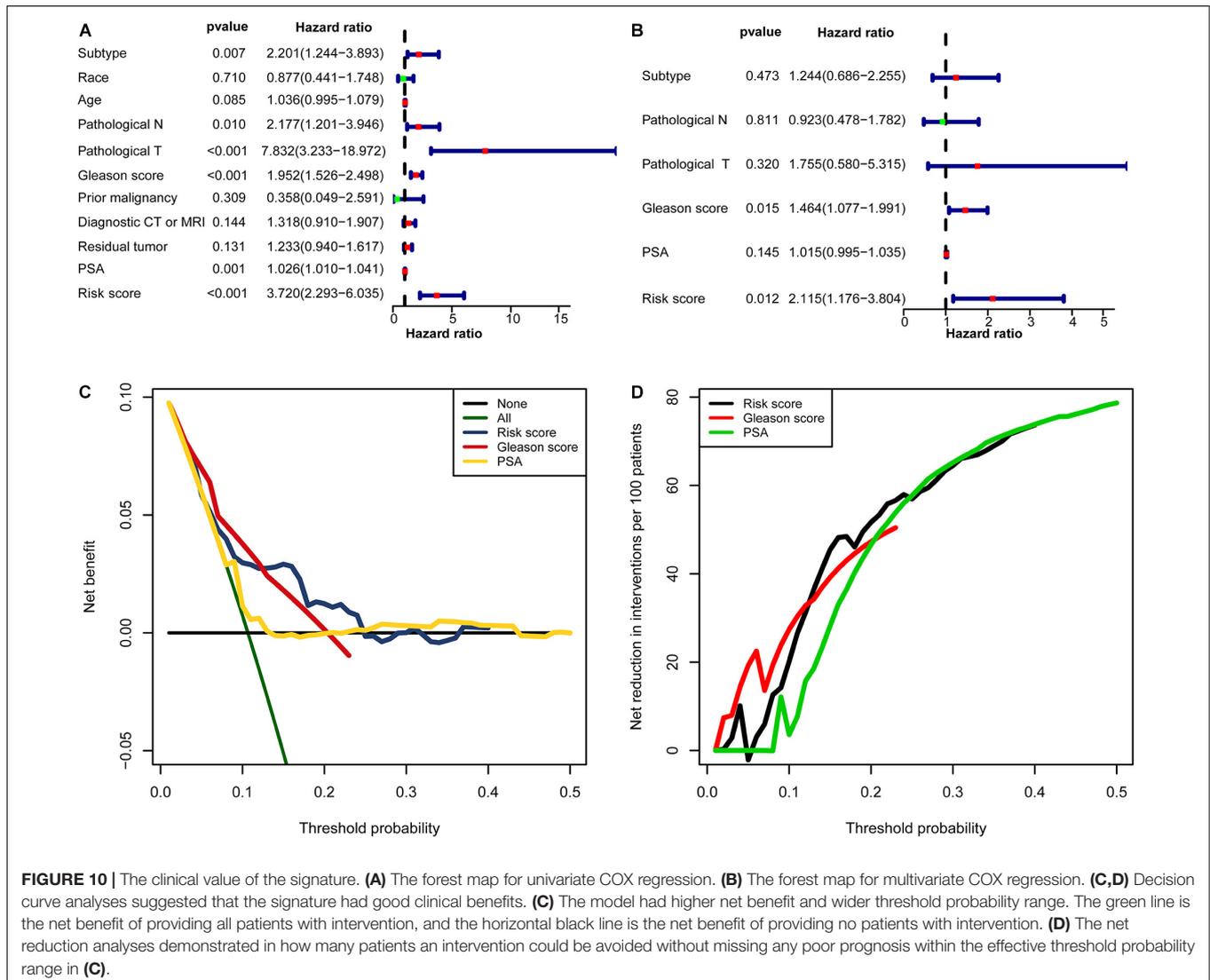
As shown in **Figure 8A**, when eight genes were included in the signature, the C-index value was maximized. Accordingly, the eight-gene signature had the best predictive value during the training process. **Figure 8B** presents the coefficients for each gene during the training process. Finally, an eight-gene signature for predicting the risk score was constructed as follows:

$$\begin{aligned} \text{Risk score} = & 0.210 \times \text{expression of TMEM132A} + 0.339 \\ & \times \text{expression of CENPF} + 0.020 \times \text{expression of} \\ & \text{DEPDC1B} + 0.081 \times \text{expression of TTK} + 0.258 \\ & \times \text{expression of CBX2} + 0.065 \times \text{expression of} \\ & \text{TOP2A} - 0.388 \times \text{expression of CDC25C} - 0.038 \\ & \times \text{expression of PARM1} \end{aligned}$$

Patients in the training and internal validation set were arranged in ascending order based on risk scores (**Figures 8C,F**). Setting the median risk score as the threshold, the frequencies of cancer-specific death or biochemical recurrence were higher in high-risk patients than in low-risk patients in the training and internal validation sets (**Figures 8D,G**). The expression modes of eight genes in the signature are displayed in **Figures 8E,H**. Furthermore, patients in the Methylation_H subtype had significantly higher risk scores than patients in the Methylation_L subtype (**Figure 8I**). Finally, we found that patients with cancer-specific death or biochemical recurrence had higher risk scores in the training set, internal validation set, and external validation sets (DKFZ2018, GSE70769, GSE116918, and MSKCC2010) (**Figures 8G,K** and **Supplementary Figure 2**).

Validation of the Signature

The areas under the curve of the tdROC (reflecting the effectiveness of a classifier) for the training set, internal validation



set, and external validation sets (DKFZ2018, GSE70769, GSE116918, and MSKCC2010) were 0.72, 0.66, 0.76, 0.76, 0.84, and 0.74, respectively (Figures 9A–F). Furthermore, a survival analysis revealed that a higher risk score was associated with a worse prognosis. Similar results were observed for all data sets evaluated (Figures 9G–L).

In univariate Cox regression analyses, six variables (subtype, pathological N, pathological T, Gleason score, PSA, and risk score) were associated with a worse prognosis (Figure 10A). In a multivariate Cox regression, Gleason score and risk score were identified as independent indicators for prognosis (Figure 10B). Within a wider range of threshold probabilities, the clinical net benefit was greater for the risk score than for the PSA level or Gleason score (Figure 10C). Within this range of threshold probabilities, the signature provides a more accurate prediction, thereby reducing the number of patients with a worse prognosis (Figure 10D).

DISCUSSION

We have recently described the advantages and necessity of multi-omics approaches for studies of PCa (Zhang et al., 2020c). In this study, we identified two subtypes with different DNA methylation statuses and found that the Methylation_H subtype was related to a worse prognosis. The subtypes were then comprehensively compared with respect to the epigenome, genome, transcriptome, disease status, immune cell infiltration, and function.

The mRNA levels of *SPOP*, which is the most frequently mutated tumor-suppressor gene in primary PCa, were lower in the Methylation_H subtype than in the Methylation_L subtype (Barbieri et al., 2012; An et al., 2014). Additionally, the mutation frequency in *SPOP* was higher in the Methylation_H subtype, and *SPOP* expression was lower in mutants. Another tumor suppressor that is downregulated early in the development of PCa, *RND3*, was expressed at significantly lower levels in the Methylation_H subtype than in the Methylation_L subtype (Bektic et al., 2005). The single-copy deletion of *RND3* was more frequent in the Methylation_H subtype and this deletion corresponded with the downregulation of *RND3*. AR scores, which reflect disease progression, were also significantly higher in the Methylation_H subtype than in the Methylation_L subtype. M1 and M2 macrophages showed a greater degree of infiltration in Methylation_H. The polarization of macrophages into the M1 and M2 phenotypes plays a pivotal role in ovarian cancer initiation, progression, and metastasis and provides targets for macrophage-centered treatment in the cancer microenvironment (Cheng et al., 2019). Consistent with these previous findings, we found that increased levels of M1 and M2 macrophages in the tumor microenvironment were related to a worse prognosis in PCa.

In the Methylation_H subtype, E2F, MYC, mTORC1, and G2M checkpoint were activated. E2F, MYC, and mTORC1 have been shown to promote the development of PCa (Huang et al., 2017; Zhang et al., 2017; Labbé et al., 2019). Furthermore, G2M checkpoint activation is related to a reduced cancer sensitivity to chemotherapy or radiation (Morgan, 2007). Furthermore,

the mRNA expression levels of *AURKA*, *DLGAP5*, *FOXD1*, *KIF4A*, *MELK*, *MYBL2*, *SPAG5*, and *TPX2* were significantly higher in the Methylation_H subtype than in the Methylation_L subtype. The gain of the *AURKA* oncogene is an important genomic change related to treatment-related neuroendocrine PCa (Akamatsu et al., 2018). Androgen-dependent PCa cells need *DLGAP5* to stabilize mitotic health and function, and the knockdown of *DLGAP5* improves the efficacy of docetaxel (Hewit et al., 2018). The knockdown of *FOXD1* and *MYBL2* would inhibit the growth of androgen-independent PCa cells (Li et al., 2020; Zhang et al., 2020b). *KIF4A* plays a significant role in the progression of castration-resistant PCa and serves as a key determinant of resistance to endocrine therapy (Cao et al., 2020). *MELK* is associated with the cell survival rate and BCR in PCa (Jurmeister et al., 2018). *SPAG5* expression is significantly associated with the clinical stage, lymph node metastasis, Gleason score, and BCR (Zhang et al., 2016). The knockdown of *TPX2* increases chromosome mis-segregation and suppresses tumor cell growth in PCa (Pan et al., 2017). These genes driving the progression of PCa were all expressed more highly in the Methylation_H subtype than in the Methylation_L subtype, further supporting the high-risk characteristics of the Methylation_H subtype.

The conserved differences uncovered the high-risk characteristics of the Methylation_H subtype. We further employed WGCNA, a common method in systems biology, to identify a key gene module; this module was related to the Gleason score, PSA, and Methylation_H subtype. Survival-associated DEGs from this gene module were used to construct an eight-gene signature for predicting risk. The effectiveness of the signature was validated in TCGA and another four public datasets (DKFZ2018, GSE70769, GSE116918, and MSKCC2010). With respect to the clinical applications of these findings, we have the following suggestions. Because RNA-seq data in TPM format were used to train the signature, we suggest employing the same data format of data in clinical applications. Considering batch effects of measurement techniques, gene expression levels should be measured by similar techniques, even though the signature performed well in the validation data sets, in which genes were profiled by array-based methods. Furthermore, the risk levels were determined by the median risk score in the patient cohorts. In the future, the study cohort should be further expanded to obtain a more objective and stable threshold range.

Collectively, we identified two subtypes with different methylation statuses at eight CpG sites and evaluated the high-risk characteristics of the Methylation_H subtype based on epigenomics, genomics, transcriptomics, disease status, immune cell infiltration, and functional analyses. Finally, based on these two novel subtypes, an eight-gene predictive signature was constructed and validated using various public datasets.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

EZ, MZ, and YS were responsible for the design and conception of the research project. EZ, YS, MZ, FS, OM, JH, YG, and HW contributed to the data acquisition or data analysis and data cleaning. EZ and YS participated in the drafting of the manuscript and the rigorous modification of the manuscript to clearly convey the research contents. All authors are responsible for the authenticity and reliability of this study and have no objection to the final submitted manuscript.

FUNDING

This research was supported by the National Natural Science Foundation of China (No. 81802540), the Natural Science Foundation of Liaoning Province (No. 20180550985), the Shenyang Science and Technology Program for Young Innovative Talents (No. RC190386), and the 345 Talent Project of Shengjing Hospital of China Medical University.

REFERENCES

- Abeshouse, A., Ahn, J., Akbani, R., Ally, A., Amin, S., Andry Christopher, D., et al. (2015). The molecular taxonomy of primary prostate cancer. *Cell* 163, 1011–1025.
- Akamatsu, S., Inoue, T., Ogawa, O., and Gleave, M. E. (2018). Clinical and molecular features of treatment-related neuroendocrine prostate cancer. *Int. J. Urol.* 25, 345–351. doi: 10.1111/iju.13526
- An, J., Wang, C., Deng, Y., Yu, L., and Huang, H. (2014). Destruction of full-length androgen receptor by wild-type SPOP, but not prostate-cancer-associated mutants. *Cell Rep.* 6, 657–669. doi: 10.1016/j.celrep.2014.01.013
- Barbieri, C. E., Baca, S. C., Lawrence, M. S., Demichelis, F., Blattner, M., Theurillat, J.-P., et al. (2012). Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat. Genet.* 44, 685–689. doi: 10.1038/ng.2279
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 41, D991–D995. doi: 10.1093/nar/gks1193
- Bektic, J., Pfeil, K., Berger, A. P., Ramoner, R., Pelzer, A., Schäfer, G., et al. (2005). Small G-protein RhoE is underexpressed in prostate cancer and induces cell cycle arrest and apoptosis. *Prostate* 64, 332–340.
- Blanche, P., and Blanche, M. P. (2013). *Package 'TimeROC': Time-Dependent ROC Curve and AUC for Censored Survival Data*. Vienna: R Foundation for Statistical Computing.
- Blum, A., Wang, P., and Zenklusen, J. C. (2018). SnapShot: TCGA-Analyzed Tumors. *Cell* 173:530. doi: 10.1016/j.cell.2018.03.059
- Cancer Genome Atlas Research Network (2015). The molecular taxonomy of primary prostate cancer. *Cell* 163, 1011–1025. doi: 10.1016/j.cell.2015.10.025
- Cao, Q., Song, Z., Ruan, H., Wang, C., Yang, X., Bao, L., et al. (2020). Targeting the KIF4A/AR Axis to reverse endocrine therapy resistance in castration-resistant prostate cancer. *Clin. Cancer Res.* 26, 1516–1528. doi: 10.1158/1078-0432.CCR-19-0396
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2, 401–404. doi: 10.1158/2159-8290.CD-12-0095
- Chang, A. J., Autio, K. A., Roach, M., and Scher, H. I. (2014). High-risk prostate cancer-classification and therapy. *Nat. Rev. Clin. Oncol.* 11, 308–323. doi: 10.1038/nrclinonc.2014.68

The 345 Talent Project includes the 30 project and the 50 project.

ACKNOWLEDGMENTS

The results shown here are in whole based upon data generated by the TCGA Research Network (<https://www.cancer.gov/tcga>), GEO (<https://www.ncbi.nlm.nih.gov/geo/>), cBioPortal (<https://www.cbioportal.org/>), and The Tumor Fusion Gene Data Portal (<https://www.tumorfusions.org/>). Furthermore, we want to thank profs. YS and MZ for their guidance and financial support for this study. At the same time, we thank the Medical Research Center of Shengjing Hospital for providing us with laboratories.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.639615/full#supplementary-material>

- Cheng, H., Wang, Z., Fu, L., and Xu, T. (2019). Macrophage polarization in the development and progression of ovarian cancers: an overview. *Front. Oncol.* 9:421. doi: 10.3389/fonc.2019.00421
- Chevillat, J. C., Karnes, R. J., Therneau, T. M., Kosari, F., Munz, J.-M., Tillmans, L., et al. (2008). Gene panel model predictive of outcome in men at high-risk of systemic progression and death from prostate cancer after radical retropubic prostatectomy. *J. Clin. Oncol.* 26, 3930–3936. doi: 10.1200/JCO.2007.15.6752
- Darnel, A. D., Lafargue, C. J., Vollmer, R. T., Corcos, J., and Bismar, T. A. (2009). TMPRSS2-ERG fusion is frequently observed in Gleason pattern 3 prostate cancer in a Canadian cohort. *Cancer Biol. Ther.* 8, 125–130.
- Daura-Oller, E., Cabre, M., Montero, M. A., Paternain, J. L., and Romeu, A. (2009). Specific gene hypomethylation and cancer: new insights into coding region feature trends. *Bioinformation* 3, 340–343.
- Engelbrechtsen, S., and Bohlin, J. (2019). Statistical predictions with glmnet. *Clin. Epigenetics* 11:123. doi: 10.1186/s13148-019-0730-1
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* 6:11. doi: 10.1126/scisignal.2004088
- Hewitt, K., Sandilands, E., Martinez, R. S., James, D., Leung, H. Y., Bryant, D. M., et al. (2018). A functional genomics screen reveals a strong synergistic effect between docetaxel and the mitotic gene DLGAP5 that is mediated by the androgen receptor. *Cell Death Dis.* 9:1069. doi: 10.1038/s41419-018-1115-7
- Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., et al. (2020). Ensembl 2021. *Nucleic Acids Res.* 49, D884–D891. doi: 10.1093/nar/gkaa942
- Huang, L., Chen, K., Cai, Z.-P., Chen, F.-C., Shen, H.-Y., Zhao, W.-H., et al. (2017). DEPDC1 promotes cell proliferation and tumor growth via activation of E2F signaling in prostate cancer. *Biochem. Biophys. Res. Commun.* 490, 707–712. doi: 10.1016/j.bbrc.2017.06.105
- Jain, S., Lyons, C. A., Walker, S. M., McQuaid, S., Hynes, S. O., Mitchell, D. M., et al. (2018). Validation of a Metastatic Assay using biopsies to improve risk stratification in patients with prostate cancer treated with radical radiation therapy. *Ann. Oncol.* 29, 215–222. doi: 10.1093/annonc/mdx637
- Jurmeister, S., Ramos-Montoya, A., Sandi, C., Pérttega-Gomes, N., Wadhwa, K., Lamb, A. D., et al. (2018). Identification of potential therapeutic targets in prostate cancer through a cross-species approach. *EMBO Mol. Med.* 10:e8274. doi: 10.15252/emmm.201708274
- Kassambara, A., Kosinski, M., Bieček, P., and Fabian, S. (2017). *Package 'survminer'. Drawing Survival Curves using 'ggplot2' (R package version 03 1)*.
- Kolde, R., and Kolde, M. R. (2015). Package 'pheatmap'. *R Package* 1, 790.

- Kuhn, M. (2008). Building predictive models in R using the caret package. *J. Stat. Softw.* 28, 1–26.
- Kuner, R., Fálth, M., Pressinotti, N. C., Brase, J. C., Puig, S. B., Metzger, J., et al. (2013). The maternal embryonic leucine zipper kinase (MELK) is upregulated in high-grade prostate cancer. *J. Mol. Med. (Berl.)* 91, 237–248. doi: 10.1007/s00109-012-0949-1
- Labbé, D. P., Zadra, G., Yang, M., Reyes, J. M., Lin, C. Y., Cacciatori, S., et al. (2019). High-fat diet fuels prostate cancer progression by rewiring the metabolome and amplifying the MYC program. *Nat. Commun.* 10:4358. doi: 10.1038/s41467-019-12298-z
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9:559. doi: 10.1186/1471-2105-9-559
- Li, G., Jiang, Y., Lyu, X., Cai, Y., Zhang, M., Wang, Z., et al. (2019). Deconvolution and network analysis of IDH-mutant lower grade glioma predict recurrence and indicate therapeutic targets. *Epigenomics* 11, 1323–1333. doi: 10.2217/epi-2019-0137
- Li, X., Jiao, M., Hu, J., Qi, M., Zhang, J., Zhao, M., et al. (2020). miR-30a inhibits androgen-independent growth of prostate cancer via targeting MYBL2, FOXD1, and SOX4. *Prostate* 80, 674–686. doi: 10.1002/pros.23979
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. doi: 10.1093/bioinformatics/btr260
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550.
- Mehra, R., Tomlins, S. A., Shen, R., Nadeem, O., Wang, L., Wei, J. T., et al. (2007). Comprehensive assessment of TMPRSS2 and ETS family gene aberrations in clinically localized prostate cancer. *Mod. Pathol.* 20, 538–544.
- Mikeska, T., and Craig, J. M. (2014). DNA methylation biomarkers: cancer and beyond. *Genes (Basel)* 5, 821–864. doi: 10.3390/genes5030821
- Monti, S., Tamayo, P., Mesirov, J., and Golub, T. (2003). Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach. Learn.* 52, 91–118. doi: 10.1023/A:1023949509487
- Morgan, D. O. (2007). *The Cell Cycle: Principles of Control*. London: New science press.
- Newman, A. M., Steen, C. B., Liu, C. L., Gentles, A. J., Chaudhuri, A. A., Scherer, F., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* 37, 773–782. doi: 10.1038/s41587-019-0114-2
- Pan, H.-W., Su, H.-H., Hsu, C.-W., Huang, G.-J., and Wu, T. T.-L. (2017). Targeted TPX2 increases chromosome missegregation and suppresses tumor cell growth in human prostate cancer. *Oncotargets Ther.* 10, 3531–3543. doi: 10.2147/OTT.S136491
- Peng, D., Ge, G., Xu, Z., Ma, Q., Shi, Y., Zhou, Y., et al. (2018). Diagnostic and prognostic biomarkers of common urological cancers based on aberrant DNA methylation. *Epigenomics* 10, 1189–1199. doi: 10.2217/epi-2018-0017
- Perner, S., Demichelis, F., Beroukhim, R., Schmidt, F. H., Mosquera, J.-M., Setlur, S., et al. (2006). TMPRSS2:ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. *Cancer Res.* 66, 8337–8341.
- Rajput, A. B., Miller, M. A., De Luca, A., Boyd, N., Leung, S., Hurtado-Coll, A., et al. (2007). Frequency of the TMPRSS2:ERG gene fusion is increased in moderate to poorly differentiated prostate cancers. *J. Clin. Pathol.* 60, 1238–1243.
- Ross-Adams, H., Lamb, A. D., Dunning, M. J., Halim, S., Lindberg, J., Massie, C. M., et al. (2015). Integration of copy number and transcriptomics provides risk stratification in prostate cancer: a discovery and validation cohort study. *EBioMedicine* 2, 1133–1144. doi: 10.1016/j.ebiom.2015.07.017
- Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA Cancer J. Clin.* 70, 7–30. doi: 10.3322/caac.21590
- Skidmore, Z. L., Wagner, A. H., Lesurf, R., Campbell, K. M., Kunisaki, J., Griffith, O. L., et al. (2016). GenVisR: genomic visualizations in R. *Bioinformatics* 32, 3012–3014. doi: 10.1093/bioinformatics/btw325
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550.
- Svane, A. M., Soerensen, M., Lund, J., Tan, Q., Jylhävä, J., Wang, Y., et al. (2018). DNA methylation and all-cause mortality in middle-aged and elderly danish twins. *Genes (Basel)* 9:78. doi: 10.3390/genes9020078
- Therneau, T. M., and Lumley, T. (2014). Package 'survival'. *Survival Anal.* 2:3.
- Tu, J. J., Rohan, S., Kao, J., Kitabayashi, N., Mathew, S., and Chen, Y.-T. (2007). Gene fusions between TMPRSS2 and ETS family genes in prostate cancer: frequency and transcript variant analysis by RT-PCR and FISH on paraffin-embedded tissues. *Mod. Pathol.* 20, 921–928.
- Van Calster, B., Wynants, L., Verbeek, J. F. M., Verbakel, J. Y., Christodoulou, E., Vickers, A. J., et al. (2018). Reporting and interpreting decision curve analysis: a guide for investigators. *Eur. Urol.* 74, 796–804. doi: 10.1016/j.eururo.2018.08.038
- Wang, Y.-P., and Lei, Q.-Y. (2018). Metabolic recoding of epigenetics in cancer. *Cancer Commun. (Lond.)* 38:25. doi: 10.1186/s40880-018-0302-3
- Yoshihara, K., Wang, Q., Torres-García, W., Zheng, S., Vegesna, R., Kim, H., et al. (2015). The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene* 34, 4845–4854. doi: 10.1038/onc.2014.406
- Yoshimoto, M., Joshua, A. M., Chilton-Macneill, S., Bayani, J., Selvarajah, S., Evans, A. J., et al. (2006). Three-color FISH analysis of TMPRSS2/ERG fusions in prostate cancer indicates that genomic microdeletion of chromosome 21 is associated with rearrangement. *Neoplasia* 8, 465–469.
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118
- Zhang, E., He, J., Zhang, H., Shan, L., Wu, H., Zhang, M., et al. (2020a). Immune-related gene-based novel subtypes to establish a model predicting the risk of prostate cancer. *Front. Genet.* 11:595657. doi: 10.3389/fgene.2020.595657
- Zhang, E., Hou, X., Hou, B., Zhang, M., and Song, Y. (2020b). A risk prediction model of DNA methylation improves prognosis evaluation and indicates gene targets in prostate cancer. *Epigenomics* 12, 333–352. doi: 10.2217/epi-2019-0349
- Zhang, E., Zhang, M., Shi, C., Sun, L., Shan, L., Zhang, H., et al. (2020c). An overview of advances in multi-omics analysis in prostate cancer. *Life Sci.* 260:118376. doi: 10.1016/j.lfs.2020.118376
- Zhang, H., Li, S., Yang, X., Qiao, B., Zhang, Z., and Xu, Y. (2016). miR-539 inhibits prostate cancer progression by directly targeting SPAG5. *J. Exp. Clin. Cancer Res.* 35:60. doi: 10.1186/s13046-016-0337-8
- Zhang, P., Wang, D., Zhao, Y., Ren, S., Gao, K., Ye, Z., et al. (2017). Intrinsic BET inhibitor resistance in SPOP-mutated prostate cancer is mediated by BET protein stabilization and AKT-mTORC1 activation. *Nat. Med.* 23, 1055–1062. doi: 10.1038/nm.4379
- Zhao, S. G., Chen, W. S., Li, H., Foye, A., Zhang, M., Sjöström, M., et al. (2020). The DNA methylation landscape of advanced prostate cancer. *Nat. Genet.* 52, 778–789. doi: 10.1038/s41588-020-0648-8
- Zou, J., Huang, R.-Y., Jiang, F.-N., Chen, D.-X., Wang, C., Han, Z.-D., et al. (2018). Overexpression of TPX2 is associated with progression and prognosis of prostate cancer. *Oncol. Lett.* 16, 2823–2832. doi: 10.3892/ol.2018.9016

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhang, Shiori, Mu, He, Ge, Wu, Zhang and Song. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.