



Expression patterns and a prognostic model of m⁶A-associated regulators in prostate adenocarcinoma

Song Ou-Yang^{1,2} , Ji-Hong Liu¹  & Qin-Zhang Wang^{*,2} 

¹Department of Urology, Tongji Hospital, Tongji Medical College, Huazhong University of Science & Technology, Wuhan, Hubei, 430030, China

²Department of Urology, First Affiliated Hospital, School of Medicine, Shihezi University, Shihezi, Xinjiang, 832008, China

*Author for correspondence: wqz1969@sina.com.cn

Aim: To study the expression patterns and prognostic value of the m⁶A-associated regulators in prostate adenocarcinoma (PRAD). **Materials & methods:** The mRNA expression and clinical data were downloaded from 'The Cancer Genome Atlas database'. The m⁶A-associated variants were downloaded from m⁶AVar database, and combined with 14 common m⁶A regulators for subsequent analysis. One-way analysis of variance, univariate Cox regression analysis and least absolute shrinkage and selection operator algorithm were successively applied to obtain the ultimate regulators and prognostic model. Finally, consensus clustering, protein–protein interaction (PPI) and enrichment analysis were performed. **Result:** Nine regulators were obtained. PRAD patients could be classified into two risk groups and subclasses with significant survival differences by the prognostic model and consensus clustering, respectively. **Conclusion:** All these nine regulators were related to prognosis in PRAD, and could be used as clinical biomarkers.

First draft submitted: 22 February 2020; Accepted for publication: 4 September 2020; Published online: 21 December 2020

Keywords: expression pattern • LASSO • m⁶A • m⁶Avar • PRAD • prognosis model • regulator • variant

Prostate adenocarcinoma (PRAD) is the second most commonly occurring cancer in men worldwide, with an estimated 1.27 million diagnoses worldwide in 2018, accounting for 13.5% of all male cancers diagnosed [1]. Therefore, it is urgent to study the occurrence, development, diagnosis and treatment of PRAD.

Recent studies [2–5] have found that epigenetics plays an important role in the occurrence and development of tumors. The diversity of RNA in cell biology processes, such as rRNA, tRNA, mRNA, snRNA and other RNA chemical modifications has received additional attention [6]. Among them, mRNA modification plays a critical role in regulating post-transcriptional levels of gene expression. In eukaryotes, N⁶-methyladenosine (m⁶A) is the most common form of mRNA modification [3], which refers to the methylation modification on the 6th nitrogen atom of adenine, and account for about 0.1–0.4% of all adenine bases [7,8]. It is reported that m⁶A exists widely in the transcriptome, with more than 7600 genes and more than 300 noncoding RNA having m⁶A modifications [9]. According to the different roles of the m⁶A regulators in the methylation process, the most common m⁶A regulators can be divided into three categories [10]: 'writers' or methyltransferases (*METTL3*, *METTL14*, *WTAP*, *KIAA1429*, *RBM15* and *ZC3H13*), 'readers' or binding proteins (*YTHDC1*, *YTHDC2*, *YTHDF1*, *YTHDF2*, *YTHDF3* and *HNRNPC*), and 'erasers' or demethylases (*FTO*, *ALKBH5*), for a total of 14 common m⁶A regulators. In addition, recent studies have also revealed that m⁶A variants are closely related to dysregulation in cellular processes, leading to serious diseases such as cancer [11,12].

However, at present, the expression pattern, mechanism and prognostic value of m⁶A-associated regulators in the development of PRAD have not been studied comprehensively. In the present paper, we used public data to study the expression patterns and construct a prognostic model of m⁶A-associated regulators in PRAD, and to explore biomarkers related to clinical prognosis.

Materials & methods

Data acquisition

The mRNA expression profiles and clinical data in PRAD patients were downloaded from ‘The Cancer Genome Atlas’ (TCGA) and The Genotype Tissue Expression (GTEx) database using the R package ‘TCGAbiolinks’ [13]. The m⁶A-associated variants related to PRAD were downloaded from the m⁶AVar database [14]. The candidate m⁶A-associated regulators were composed of the 14 common m⁶A RNA methylation regulators (*METTL3*, *METTL14*, *WTAP*, *KIAA1429*, *RBM15*, *ZC3H13*, *YTHDC1*, *YTHDC2*, *YTHDF1*, *YTHDF2*, *YTHDF3*, *HNRNPC*, *FTO* and *ALKBH5*) and the m⁶A-associated variants.

Acquisition of m⁶A-associated regulators & prognostic value model

First, the initial candidate m⁶A-associated regulators, both containing the gene expression profiles and clinical information in the TCGA-PRAD dataset, were obtained. Then, one-way analysis of variance (ANOVA) was applied to screen out the m⁶A regulators related to the Gleason grade classification (using the parameters of $p < 0.05$). Univariate Cox regression analysis was performed to screen out the m⁶A-associated regulators related to the prognosis (using the parameters of $p < 0.0001$ and hazard ratio [HR] > 1). Finally, a prognostic risk-scoring value model was constructed by the least absolute shrinkage and selection operator (LASSO) algorithm using the R package ‘glmnet’ [15], then the ultimate m⁶A-associated regulators and risk scores for the PRAD patients were obtained.

The risk score = $\sum_{j=1}^n \text{Coef}_j \times X_j$ (X_j represents expression value of each gene, Coef_j represents coefficient obtained from LASSO algorithm).

Validation of the mRNA expression level & survival analysis

To observe the mRNA expression levels and survival analysis of the ultimate m⁶A-associated regulators in PRAD patients, the mRNA gene expression profiles of the ultimate regulators in PRAD and normal counterparts from TCGA target GTEx databases, and survival data in TCGA-PRAD dataset were downloaded from using the R package ‘TCGAbiolinks’ [13]. Then the R package ‘ggpubr’ was performed for statistical analysis, the R package ‘survival’ and ‘survminer’ were performed for survival analysis.

Validation of the protein expression level

To validate the trend at protein expression level of the up-regulated m⁶A-associated regulators between PRAD and normal prostate tissues, the immunohistochemistry images retrieved from the Human Protein Atlas (HPA) database [16] were investigated.

Consensus clustering & principal component analysis

To further verify the stability and accuracy, and study the function of the ultimate m⁶A-associated regulators in PRAD, a consistent clustering analysis was performed using the R package ‘ConsensusClusterPlus’ [17], and then the PRAD patients were classified into different subclasses. Next a principal component analysis (PCA) was performed to judge the correctness of the classification. The parameters were set to resample rate of 80%, maximum clustering number of ten, clustering distance of Euclidean distance and clustering method of k-means.

PPI networks analysis

To further explore the interaction networks between the ultimate regulators, we extracted the protein–protein association networks from String database [18]. Meanwhile, a Pearson correlation coefficient analysis was performed to calculate the correlation between the ultimate m⁶A-associated regulators.

Enrichment analysis

To explore the functional and signaling pathway enrichment analysis between the subclasses, the differential expression genes (DEGs) that were significantly expressed between the subclasses were first screened out using the R package ‘edgeR’ [19] (using the parameters of $p < 0.05$ and $FC \geq 1.5$). The gene ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were then carried out using the R package ‘clusterProfiler’ [20].

Table 1. The m⁶A-associated regulators related to recurrence-free survival in prostate adenocarcinoma.

Gene	p-value	HR	Lower 95% CI	Upper 95% CI
<i>ANKLE1</i>	2.03e-05	3.629255	2.006137	6.565599
<i>UIMC1</i>	7.27e-06	4.115421	2.217931	7.636257
<i>CKAP2L</i>	2.51e-05	3.299519	1.89379	5.748695
<i>DNAH17</i>	8.64e-05	3.359837	1.83472	6.152713
<i>C19orf57</i>	2.22e-06	3.959251	2.239214	7.000523
<i>RFC5</i>	2.82e-05	3.492185	1.944895	6.270446
<i>COL11A2</i>	3.00e-05	3.279385	1.877372	5.728415
<i>WDR90</i>	1.64e-05	3.693391	2.0385	6.691752
<i>AMDHD2</i>	4.97e-05	3.489443	1.907865	6.382115
<i>DGKQ</i>	9.17e-05	3.150163	1.772682	5.598028
<i>DNHD1</i>	9.85e-06	3.602239	2.040771	6.358443

The HR, 95% CI were calculated by uni-variate Cox regression analysis.
HR: Hazard ratio; RFS: Recurrence-free survival.

Statistical analysis

All statistical analyses were completed in R software version 3.6.2, using RStudio version 1.2.5001. Aov function was used for one-way ANOVA, R package 'pheatmap' for cluster analysis, coxph function for univariate Cox regression analysis, R package 'glmnet' [15] for LASSO algorithm, R package 'survival' and 'survminer' for survival analysis, R package 'corrplot' for Pearson correlation coefficient analysis, prcomp function for PCA analysis. Categorical variables were compared using χ^2 analysis, Fisher's exact or the binomial tests of proportions. Continuous data were compared using independent *t*-tests. Kaplan–Meier analysis and Cox proportional hazard models were applied for survival analysis. Multivariate Cox proportional hazards regression model was adopted to adjust for covariate effects, and stratification analysis was used to reduce the potential confounding effect on the estimation of HR. Missing data were coded and excluded from the analysis. $p < 0.05$ was considered to indicate a statistically significant difference.

Results

Acquisition of the initial m⁶A regulators related to Gleason grade in PRAD

First, a candidate gene set containing 913 m⁶A-associated methylation regulators was obtained. Among them, 889 m⁶A-associated variants related to PRAD were obtained from the m⁶aVar database. Next, a total of 898 m⁶A regulators and 368 cases of PRAD patients with both mRNA expression profiles and clinical information in TCGA-PRAD dataset were acquired.

According to the Gleason grading system, the PRAD patients were divided into three grades: low grade (Gleason score ≤ 6), medium grade (Gleason score = 7) and high grade (Gleason score = 8–10), of which 41 cases of low grade, 186 cases of medium grade and 141 cases of high grade, respectively. Finally, 329 m⁶A-associated regulators significantly related to Gleason grade classification were initially obtained by one-way ANOVA ($p < 0.05$).

Acquisition of the m⁶A regulators related to prognosis in PRAD

The PRAD patients were classified into high- and low-expression groups according to the mean value of gene expression values. Finally, 11 m⁶A regulators related to the recurrence-free survival (RFS) in patients with PRAD were screened out (Table 1; $p < 0.0001$ and HR > 1) by univariate Cox regression analysis on those initial m⁶A regulators. The overall survival analysis was not considered as only eight cases of deaths were found during present study. The result suggested that these regulators might be primary prognostic risk factors in PRAD.

Construction of the prognostic risk-scoring model in PRAD

To better predict the effect of the 11 m⁶A-associated regulators on the clinical prognostic outcomes in PRAD patients, a LASSO regression model algorithm was performed to construct a prognostic risk-scoring model (Figure 1A). According to the minimum criteria of the prognostic model, nine m⁶A-associated regulators were ultimately obtained. The coefficients obtained from LASSO algorithm were used to further calculate the risk score (Figure 1B). Then divided the PRAD patients into high- and low-risk groups based on the mean risk score, of which 163 cases of patients in high-risk group and 205 cases of patients in low-risk group. At last, a Kaplan–Meier RFS survival

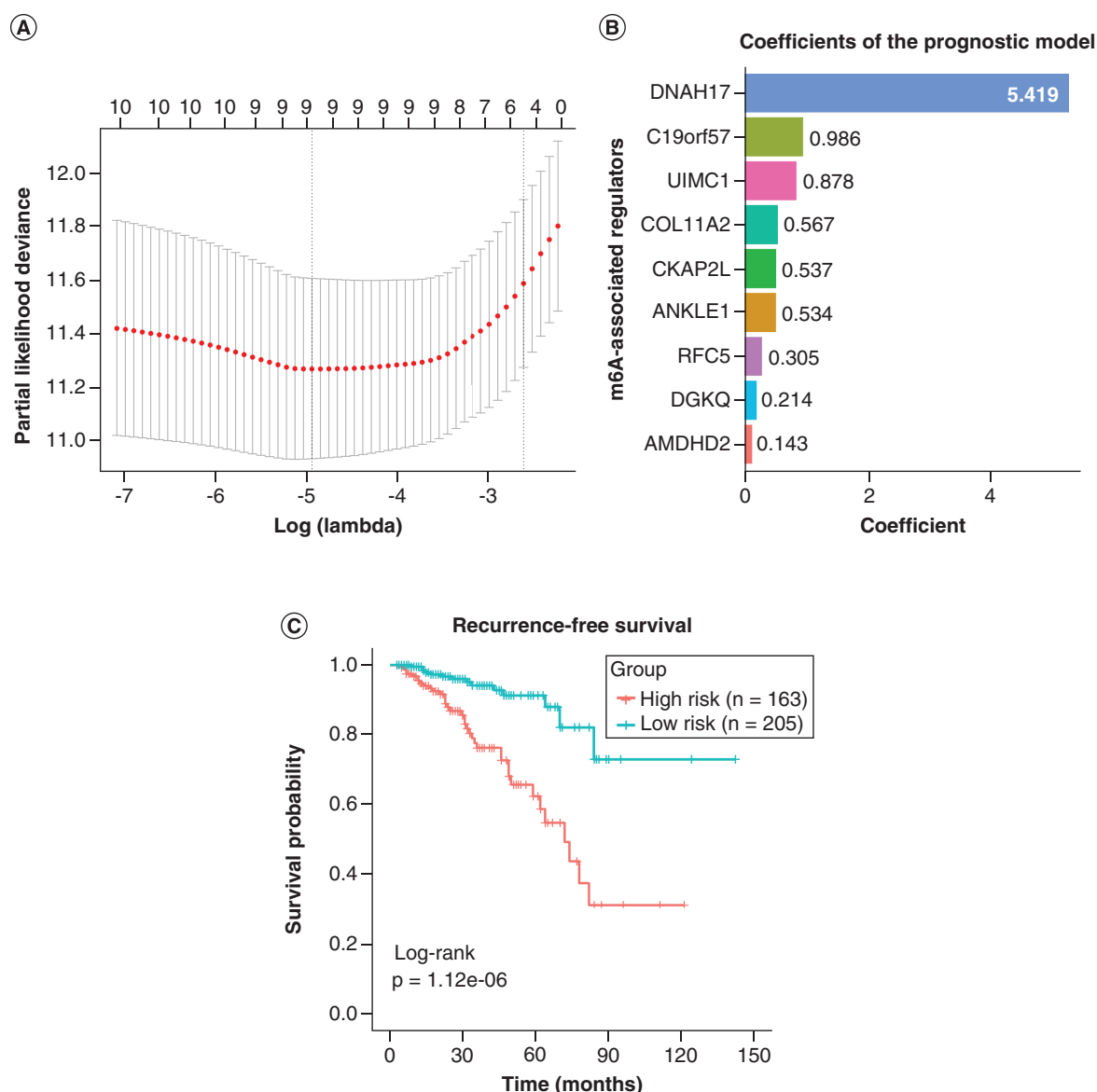


Figure 1. Construction of the prognostic model in prostate adenocarcinoma. (A) Risk signature with the 11 m⁶A-associated regulators. Nine m⁶A regulators were ultimately obtained based on the value of lambda. **(B)** The risk coefficients of the ultimate m⁶A-associated regulators. The coefficients were calculated by multi-variate Cox regression using LASSO algorithm. **(C)** Kaplan–Meier RFS curves for TCGA-PRAD dataset assigned to high- and low-risk groups based on the risk score ($p = 1.12e-06$).

LASSO: Least absolute shrinkage and selection operator; PRAD: Prostate adenocarcinoma; RFS: Recurrence-free survival; TCGA: The Cancer Genome Atlas.

analysis based on the risk score was performed to verify the model, and the result indicated significant difference between the two groups ($p = 1.12e-06$, Figure 1C).

Comparison of the clinicopathological characteristics in PRAD between high- & low-risk groups

Next, we compared the clinicopathological characteristics between the two risk groups (Table 2). The results showed significant differences in age, Gleason grade, stage T and stage N. The age of high-risk group was younger than low-risk group, and most of them were high grade, stage T2 and T3, and stage N1. The low-risk group were mostly medium and low grade, stage T1, and stage N0 (Figure 2A & B).

Table 2. Comparison of the clinicopathological characteristics between high- and low-risk groups.

Parameters	Group	High risk (n = 163)	Low risk (n = 205)	χ^2 -value	p-value
Age (year)	>60	63	107	6.168	0.013
	≤60	100	98		
Gleason grade	High grade	103	38	77.169	2.20e-16
	Medium grade	47	139		
	Low grade	13	28		
Stage T	T1	44	84	-	0.008684 [†]
	T2	70	75		
	T3	26	21		
	T4	2	0		
	TX	21	25		
Stage N	N0	112	151	16.924	3.89e-05
	N1	42	15		
	Nx	9	39		
Stage M	M0	150	189	-	-
	M1	0	0		
	Mx	13	16		

[†] Represents Fisher's exact test.

The mRNA expression levels & survival analysis in PRAD

The mRNA expression profiles of the ultimate regulators between PRAD and normal counterparts were examined in TCGA target GTEx database, and the results indicated that all were significantly expressed in PRAD patients, of which three regulators (*AMDHD2*, *CKAP2L* and *RFC5*) were significantly up-regulated, and six regulators (*ANKLE1*, *C19orf57*, *COL11A2*, *DGKQ*, *DNAH17* and *UIMC1*) significantly down-regulated (Figure 3A). Next, we performed a Kaplan–Meier progression-free survival (PFS) analysis on the different expression of these ultimate regulators, and the results also showed significantly significant differences (Figure 3B). Finally, multiple risk factors for PRAD were analyzed by univariate and multivariate Cox hazard models, the results of univariate analysis showed that the survival prognosis of PRAD is associated with Gleason grade, stage T, stage N and all the nine ultimate regulators expression, the present results also revealed the expression of nine ultimate regulators was associated with PFS prognosis in multivariate analysis (Table 4).

The protein expression levels of the up-regulated regulators in PRAD

We next validated the trend at the protein level of the three up-regulated regulators (*AMDHD2*, *CKAP2L* and *RFC5*) between the glandular cells (normal prostate tissue) and PRAD (tumor tissue). In the immunohistochemistry data from the HPA database, the PRAD patients' samples for the three up-regulated regulators had moderate or weak staining signals, whereas no staining was detected in normal prostate tissue (Figure 4A–F).

Consensus clustering of m⁶A-associated regulators identified two subclasses of PRAD

To study the function of the ultimate regulators in PRAD, a consensus clustering analysis was performed on the 368 cases of PRAD patients according to the expression similarity of the nine ultimate m⁶A-associated regulators. The results indicated that during the process from $k = 2$ to $k = 10$, with the increase of clustering stability, $k = 2$ was a suitable classification number (Figure 5A–C). Therefore, the PRAD patients could be divided into two clinical subclasses, of which 177 cases of patients in class 1 and 191 cases of patients in class 2. Finally, we analyzed the two subclasses by PCA, and the results showed class 1 and class 2 could gathered together, respectively (Figure 5D). These results indicate that the results of our classification by m⁶A-associated regulators were correct.

Comparison of the clinicopathological characteristics & prognosis between the two subclasses

Next, we compared the clinicopathological characteristics and RFS prognosis between the two subclasses. The results indicated significant differences in Gleason grade, stage T stage and stage N. In class 2, high grade, stage T2 and T3, and stage N1 were most common, while in class 1 were mostly medium and low grade, stage T1, and stage N0 (Table 3 & Figure 6A). The results of Kaplan–Meier RFS survival analysis based on the consensus clustering

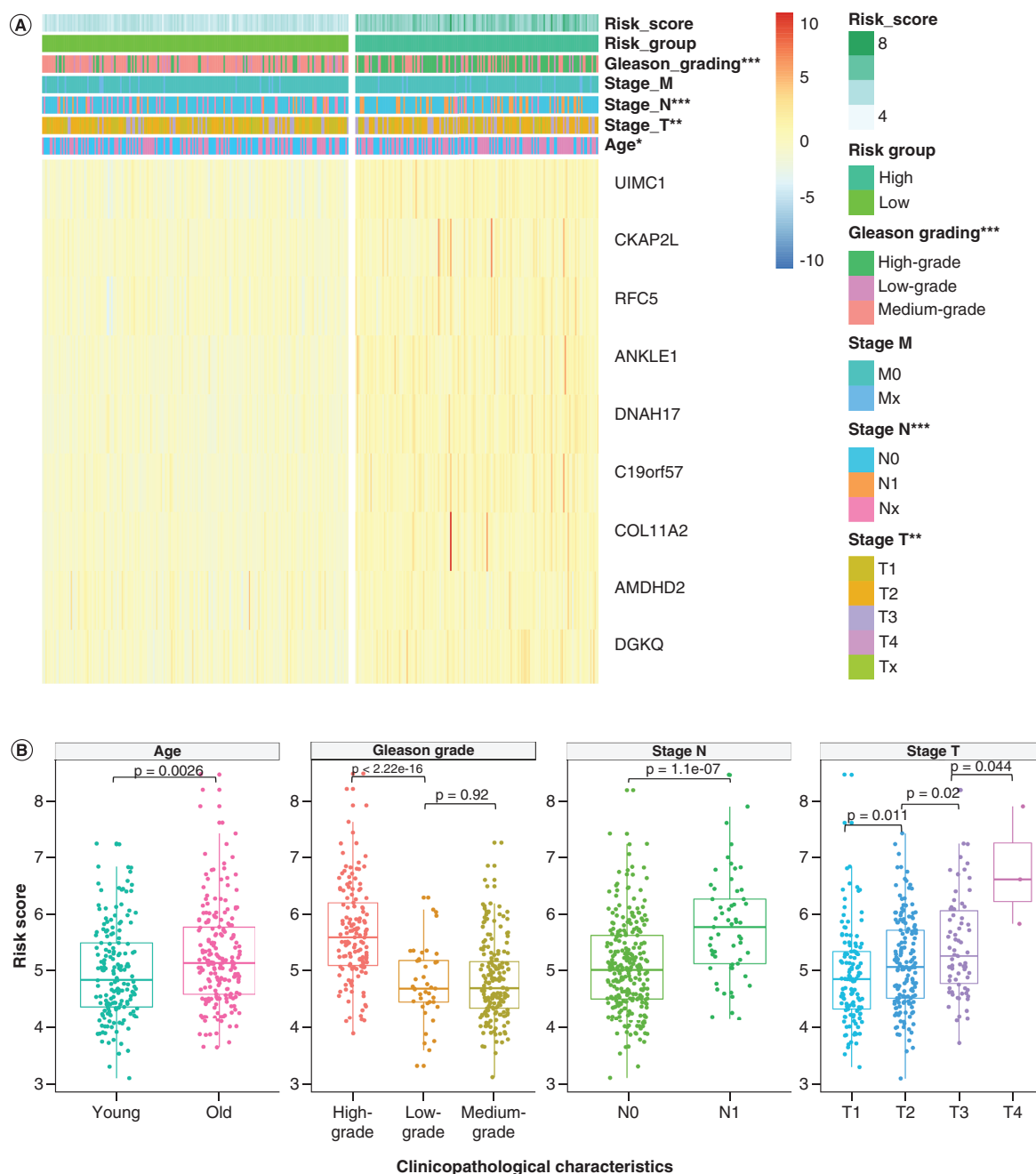


Figure 2. Clinicopathological characteristics in prostate adenocarcinoma patients between high- and low-risk groups. (A) The heatmap showed the expression levels of the nine m⁶A-associated regulators in high- and low-risk groups. The distribution of clinicopathological features was compared between the high- and low-risk groups. **(B)** Distribution of risk scores in TCGA-PRAD dataset stratified by age, Gleason grade, stage T and stage N. Note: Tx, Nx, Mx were not analyzed. *p < 0.01; **p < 0.001; ***p < 0.001. PRAD: Prostate adenocarcinoma; TCGA: The Cancer Genome Atlas.

analysis also suggested that class 2 was significantly lower than class 1 (Figure 6B). In conclusion, we believed that class 2 was the high-risk group.

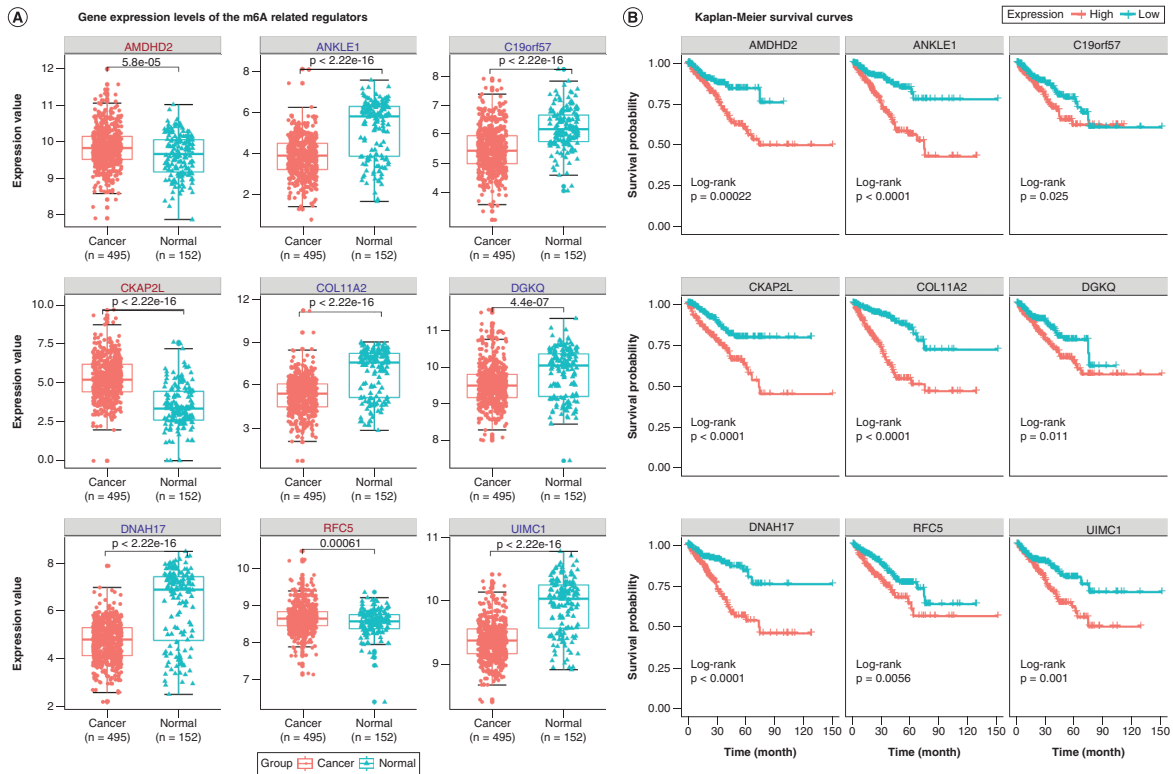


Figure 3. Analysis of the mRNA expression levels and survival in prostate adenocarcinoma patients. (A) All the nine regulators were significantly expressed in patients with PRAD compared with the normal counterparts, of which three regulators (*AMDHD2*, *CKAP2L* and *RFC5*) were significantly up-regulated, and six regulators (*ANKLE1*, *C19orf57*, *COL11A2*, *DGKQ*, *DNAH17* and *UIMC1*) were significantly down-regulated. **(B)** Kaplan–Meier PFS curves for TCGA-PRAD dataset assigned to high (red) and low (green) expression on the nine regulators showed all significantly different.

PFS: Progression-free survival; PRAD: Prostate adenocarcinoma; TCGA: The Cancer Genome Atlas.

Characteristic	Group	Class 1 (n = 177)	Class 2 (n = 191)	χ ² -value	p-value
Age (years)	>60	88	82	1.4398	0.2302
	≤60	89	109		
Gleason grade	High grade	36	105	47.39	5.12e-11
	Medium grade	113	73		
	Low grade	28	13		
T stage	T1	77	51	–	0.0001956†
	T2	63	82		
	T3	14	33		
	T4	0	2		
	TX	23	23		
N stage	N0	132	131	14.91	0.000113
	N1	12	45		
	Nx	33	15		
M stage	M0	163	176	–	–
	M1	0	0		
	Mx	14	15		

† Represents Fisher's exact test.

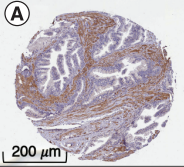
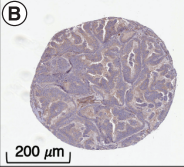
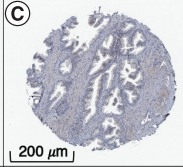
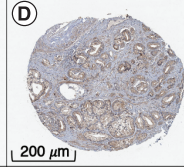
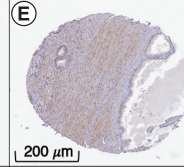
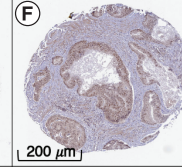
					
Prostate	Prostate cancer	Prostate	Prostate cancer	Prostate	Prostate cancer
HPA041184 Male, age 76 Prostate (T-77100) Normal tissue, NOS (M-00100) Patient id: 2932	HPA041184 Male, age 61 Prostate (T-77100) Adenocarcinoma, High grade (M-814033) Patient id: 3457	HPA040057 Male, age 76 Prostate (T-77100) Normal tissue, NOS (M-00100) Patient id: 2932	HPA040057 Male, age 75 Prostate (T-77100) Adenocarcinoma, High grade (M-814033) Patient id: 3972	HPA041037 Male, age 76 Prostate (T-77100) Normal tissue, NOS (M-00100) Patient id: 2932	HPA041037 Male, age 70 Prostate (T-77100) Adenocarcinoma, High grade (M-814033) Patient id: 3563
Glandular cells	Tumor cells	Glandular cells	Tumor cells	Glandular cells	Tumor cells
Staining: Not detected Intensity: Negative Quantity: None Location: None	Staining: Low Intensity: Weak Quantity: >75% Location: Cytoplasmic/membranous	Staining: Not detected Intensity: Weak Quantity: <25% Location: Cytoplasmic/membranous	Staining: Medium Intensity: Moderate Quantity: >75% Location: Cytoplasmic/membranous	Staining: Not detected Intensity: Weak Quantity: <25% Location: Nuclear	Staining: Medium Intensity: Moderate Quantity: >75% Location: Nuclear
AMDHD2		CKAP2L		RFC5	

Figure 4. The protein expression levels of the up-regulated regulators in prostate adenocarcinoma patients. The representative protein expression of *AMDHD2* in normal prostate tissue. (B) The representative protein expression of *AMDHD2* in PRAD tissue. (C) The representative protein expression of *CKAP2L* in normal prostate tissue. (D) The representative protein expression of *CKAP2L* in PRAD tissue. (E) The representative protein expression of *RFC5* in normal prostate tissue. (F) The representative protein expression of *RFC5* in PRAD tissue. PRAD: Prostate adenocarcinoma.

Table 4. Univariate and stepwise multivariate Cox hazard analysis of risk factors.

Characteristics	Univariate analysis			Multivariate analysis		
	HR	95% CI	p-value	HR	95% CI	p-value
AMDHD2	1.32	0.96–1.83	0.008	2.85	0.03–23.53	0.046
ANKLE1	1.79	1.43–2.23	3.52e-07	0.06	0–7.01	0.025
C19orf57	2.11	1.61–2.77	7.012e-08	5.49	0.15–19.24	0.001
CKAP2L	1.49	1.28–1.74	2.98e-07	0.06	0.02–0.18	3.18e-07
COL11A2	1.6	1.39–1.85	1.70e-10	5.76	0.04–31.27	0.035
DGKQ	1.97	1.4–2.76	8.79e-05	11.7	0.23–59.96	0.013
DNAH17	2	1.53–2.62	5.30e-07	0.58	0–0.65	0.043
UIMC1	4.78	2.48–9.21	2.98e-06	0.01	0–1.9	0.018
RFC5	3.38	2.06–5.52	1.29e-06	23.88	14.87–38.6	3.88e-08
Age (years)	1.02	0.99–1.05	0.155	–	–	–
Gleason grade	4.13	2.69–6.35	9.27e-11	0.66	0–157.48	0.028
Stage T	2.69	1.79–4.03	1.68e-06	1.01	0.01–18.18	0.097
Stage M	3.58	0.49–25.82	0.207	–	–	–
Stage N	1.89	1.17–3.05	0.01	15.92	0.12–21.17	0.169

HR: Hazard ratio.

PPI analysis on the ultimate m⁶A-associated regulators

We extracted its protein–protein association networks from String’s website, and a network of nine nodes with one edge of protein interaction diagram were finally collected, in which node *ANKLE1* and *UIMC* had an interactive regulatory relationship (Figure 7A). Meanwhile, we analyzed the correlation between these ultimate m⁶A-associated regulators using Pearson correlation coefficient analysis, and the results showed that *AMDHD2* was significantly correlated with *UIMC1*, *CKAP2L* and *RFC5*, *DNAH17* was significantly correlated with *CKAP2L*, respectively (Figure 7B).

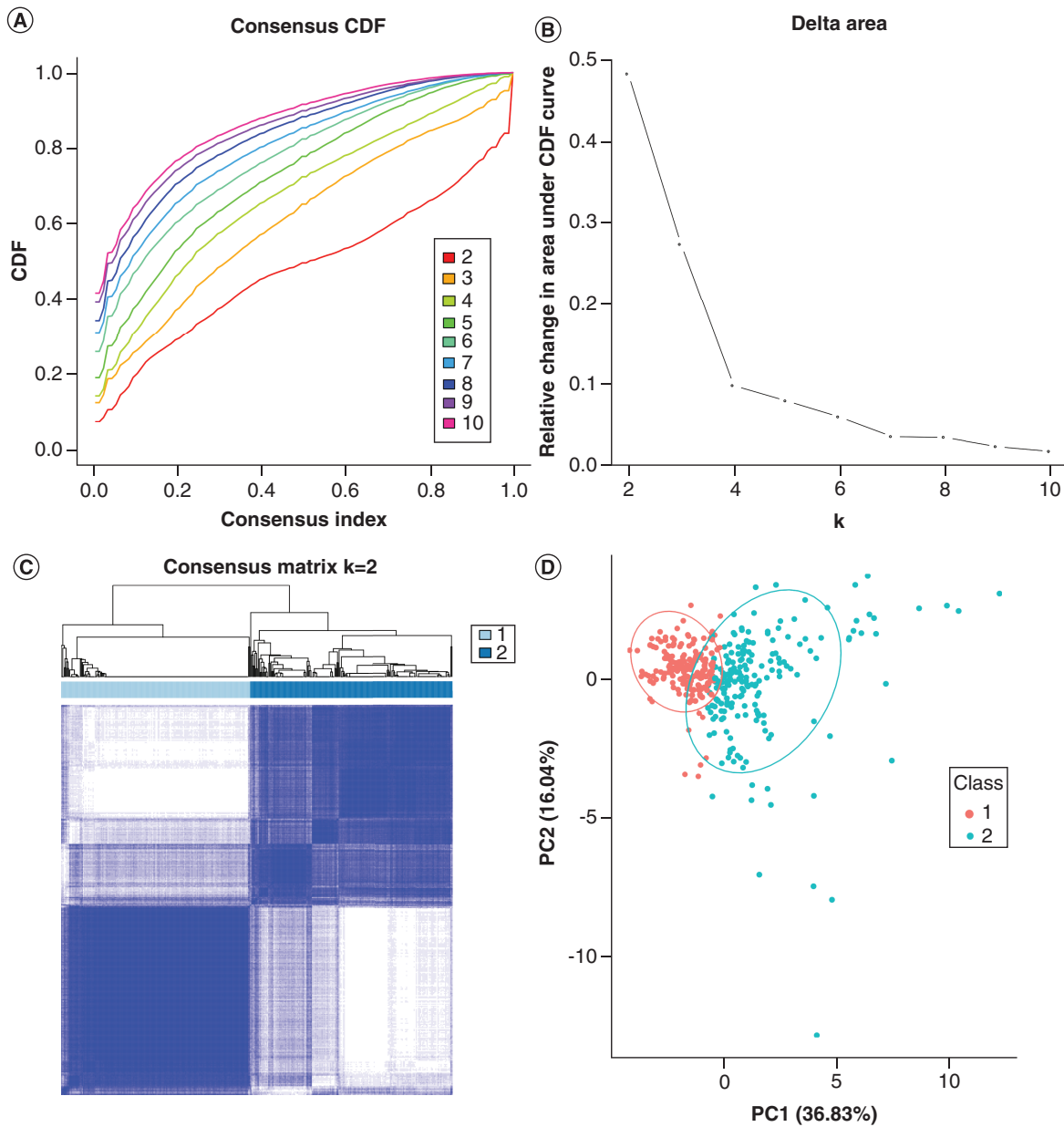


Figure 5. Identification of consensus clusters by m⁶A-associated regulators. (A) Consensus clustering CDF for k = 2–10. **(B)** Relative change in area under CDF curve from k = 2 to 10. **(C)** Consensus matrix for k = 2. **(D)** Principal component analysis of the total RNA expression profile in the TCGA dataset. PRAD in the class 1 subgroup are marked with red, and the class 2 subgroup are marked with green. CDF: Cumulative distribution function; PRAD: Prostate adenocarcinoma; TCGA: The Cancer Genome Atlas.

Enrichment analysis between the subclasses in PRAD

Next, an edgeR algorithm was conducted to obtain the DEGs that were significantly different in class 2 compared with class 1. A total of 786 DEGs were finally obtained, of which 439 DEGs were significantly up-regulated and 347 DEGs significantly down-regulated.

The GO function and KEGG signaling pathway enrichment analysis were performed on the up- and down-regulated DEGs, respectively. The GO functional annotations enrichment analysis was performed to analyze the functions in biological processes (BP), cellular components (CC) and molecular functions (MF). The up-regulated genes were mainly related to ‘microtubule cytoskeleton organization’, ‘chromosome segregation’ and ‘sister chromatid segregation’ in BP, ‘condensed chromosome’, ‘centromeric region’ and ‘condensed chromosome’

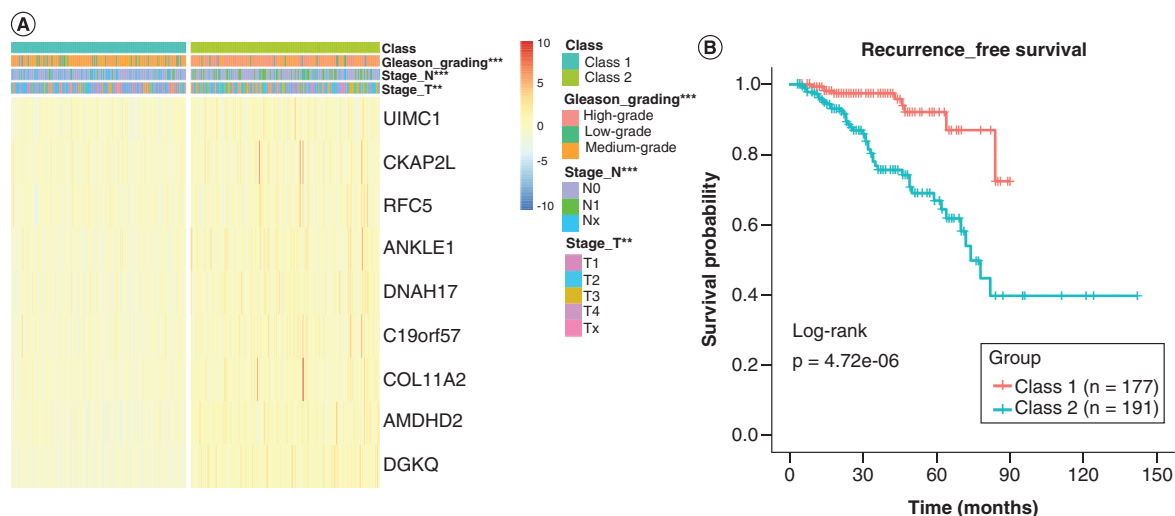


Figure 6. Clinicopathological characteristics and RFS in prostate adenocarcinoma patients between two subclasses. (A) The heatmap showed the expression levels of the nine ultimate m⁶A-associated regulators in high- and low-risk subclasses. The distribution of clinicopathological features was compared between the high- and low-risk subclasses. (B) Kaplan–Meier RFS curves for TCGA-PRAD dataset assigned to class 1 and class 2 based on the consensus clustering analysis ($p = 4.72e-06$). * $p < 0.01$; ** $p < 0.001$; *** $p < 0.001$. PRAD: Prostate adenocarcinoma; RFS: Recurrence-free survival; TCGA: The Cancer Genome Atlas.

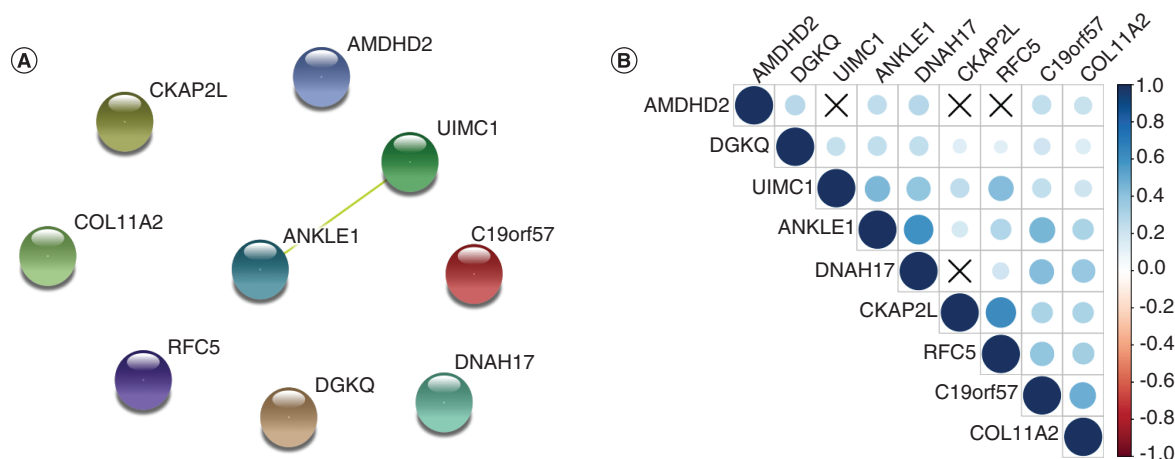


Figure 7. Protein–protein interaction analysis on the ultimate m⁶A-associated regulators. (A) Protein–protein interaction networks from String’s websites. (B) Pearson correlation analysis of the nine ultimate m⁶A-associated regulators in PRAD. ‘x’ indicates $p > 0.05$, the darker blue indicates a greater positive correlation, and the darker red indicates a greater negative correlation. PRAD: Prostate adenocarcinoma.

in CC, ‘peptidase inhibitor activity’, ‘endopeptidase inhibitor activity’ and ‘endopeptidase regulator activity’ in MF (Figure 8A). The down-regulated genes were mainly related to ‘muscle contraction’, ‘muscle system process’ and ‘heart contraction’ in BP, ‘contractile fiber’, ‘myofibril’ and ‘contractile fiber part’ in CC, ‘ion channel binding’, ‘actin binding’ and ‘structural constituent of muscle’ in MF (Figure 8B).

Furthermore, the KEGG signaling pathways analysis for up-regulated genes were mainly related to ‘cell cycle’, ‘neuroactive ligand–receptor interaction’, ‘p53 signaling pathway’ ‘endocrine resistance’ and ‘oocyte meiosis’ (Figure 8C), and the KEGG pathways analysis for down-regulated genes were mainly related to ‘vascular smooth muscle contraction’, ‘calcium signaling pathway’, ‘cGMP-PKG signaling pathway’, ‘adrenergic signaling in cardiomyocytes’ and ‘cardiac muscle contraction’.

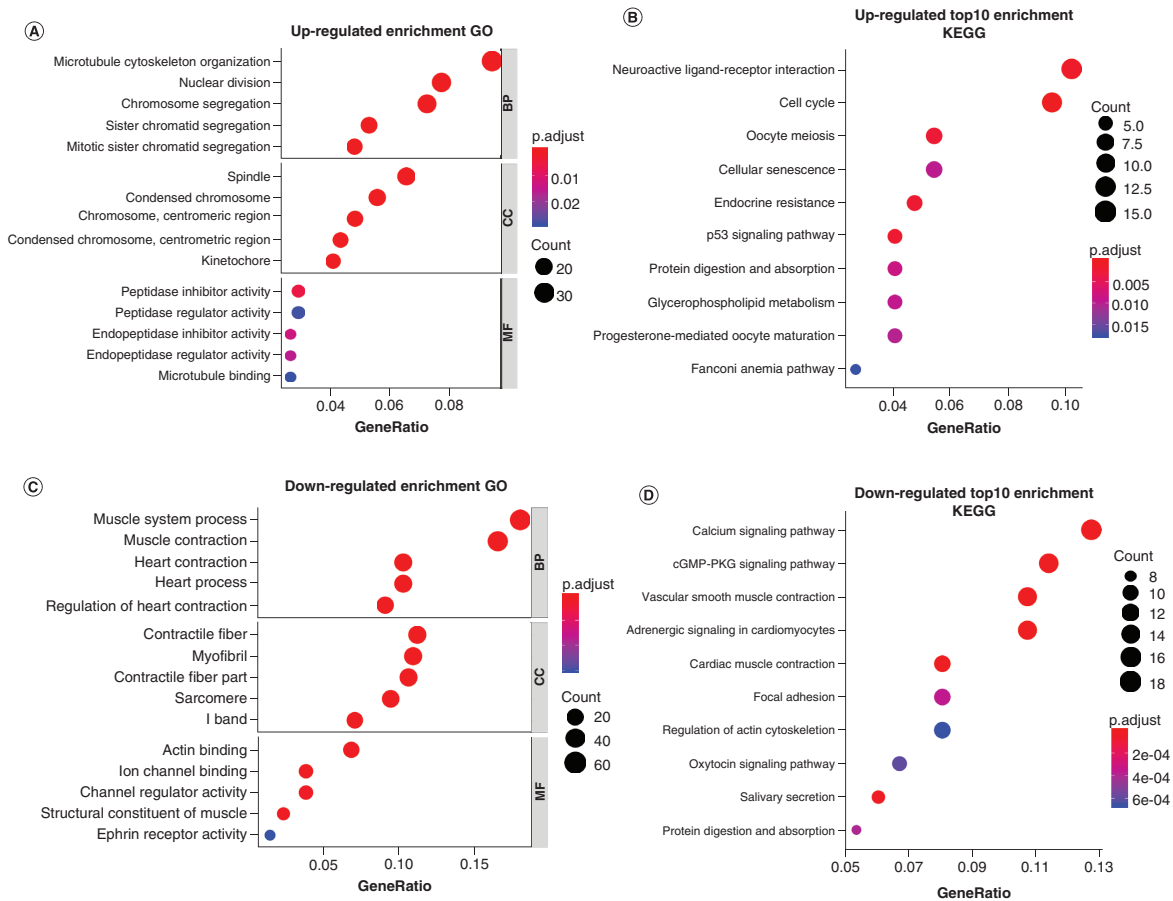


Figure 8. Dotplot for functional and signaling pathway enrichment analysis between the subclasses. (A) The GO enrichment analysis on the up-regulated DEGs. **(B)** Top 10 KEGG enrichment analysis on the up-regulated DEGs. **(C)** The GO enrichment analysis on the down-regulated DEGs. **(D)** Top 10 KEGG enrichment analysis on the down-regulated DEGs. BP: Biological process; CC: Cell component; DEG: Differentially expressed gene; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Gene and Genome; MF: Molecular function.

Discussion

The RNA modification is a post-transcriptional regulation, and more than 150 RNA modifications have been identified [8]. They are widely distributed in various types of RNA [21]. Among them, the mRNA methylation modification accounts for more than 60% [22], and N⁶-methyladenosine is the most common modification for mRNA and lncRNAs in higher organisms [8]. The m⁶A modification has also been found to be present in microRNA, circRNA, rRNA, tRNA and snoRNA [23]. The function of the m⁶A modification is determined by ‘writer’, ‘eraser’ and ‘reader’. The ‘writer’, also called RNA methyltransferases, catalyze the methylation process. However, the ‘eraser’ or demethylases can reverse the RNA methylation. In addition, the m⁶A methylation can be recognized by the ‘reader’, also called m⁶A-binding proteins. All three effectors jointly regulate the process of RNA methylation modification. As the most abundant internal modification in eukaryotic, m⁶A has been shown to play essential roles in various normal bioprocesses recently [8]. It is known that cancer has many potential links with m⁶A modifications [7], evidence also indicated that m⁶A modification and its regulators play critical roles in various cancers at level of mRNA methylation [7,22]. Besides, m⁶A is also considered to influence the lncRNA splicing [24] and miRNA processing [25], which might alter the cancer progression [26,27].

Furthermore, more studies have revealed that the m⁶A-associated variants are also closely linked to the dysregulation in cellular processes, even leading to serious diseases, such as cancer [11,12]. Functional variants, especially cancer mutations, can significantly alter the status of m⁶A, leading to the gain or loss of N⁶-methyladenosine. Meng *et al.* [28] have indicated that genetic variants in m⁶A modification genes might be promising predictors of

colorectal cancer risk. The m⁶AVar database is specifically designed to collect such functional variants, and aimed at providing potential help for revealing the functional roles of the m⁶A variants. So far, 352,014 germline mutations from dbSNP and 62,227 somatic mutations from TCGA database have been included in m⁶AVar database. In the present study, the candidate m⁶A-associated regulators were composed of the common m⁶A regulators and the m⁶A-associated variants related to PRAD. The m⁶A-associated variants related to PRAD were downloaded from the m⁶AVar database, and combined with the 14 common m⁶A regulators for subsequent analyses.

To our knowledge, the expression pattern, mechanism, clinical characteristics and prognostic value of the m⁶A-associated regulators in PRAD have been rarely reported. At the mRNA level, several studies [29–31] have indicated that *METTL3* promotes the growth and progression of PRAD. A comprehensive silico analysis found that the survival benefits of PRAD patients with overexpression of ‘reader’ and ‘writer’ in mRNA methylation was poor. In addition, at the noncoding RNA level, Daniela *et al.* [32] investigated m⁶A biological function for lncRNAs in PRAD by ELISA and western blot methods, and speculated a novel regulatory mechanism whereby VIRMA knockdown reduces the stability of oncogenic lncRNA by reducing m⁶A mark, thereby alleviating the malignant biology of prostate cancer. The potential role of m⁶A modification in PRAD was the focus on the present study. In the present paper, we systematically used bioinformatics analysis on TCGA databases, and for the first time demonstrated the expression patterns and constructed a prognostic model of the m⁶A-associated regulators in PRAD, as well as exploring their use as clinical biomarkers. However, the detailed molecular mechanisms of these biomarkers in progression of PRAD need to be explored in further studies.

In the present study, a one-way ANOVA related to Gleason grade, univariate Cox regression analysis related to RFS, and LASSO regression model algorithm in PRAD were successively applied. Finally, a prognostic risk-scoring model with nine m⁶A-associated regulators was ultimately built. Since only eight cases of death were found in the TCGA-PRAD dataset, the overall survival analysis was not considered. Following, according to the risk-scoring model, the PRAD patients were divided into high- and low-risk groups, the results of clinicopathological characteristics and RFS survival analysis both confirmed the significant difference between the two risk groups. Preliminary conclusions are that these m⁶A-associated regulators have clinical prognostic significance.

In addition, the results of mRNA expression levels and PFS analysis on the ultimate m⁶A-associated regulators in PRAD patients indicated that all these nine m⁶A-associated regulators were significantly different in PRAD patients. The results were also validated by univariate and stepwise multivariate Cox hazard analysis. Among the mRNA expression levels, three regulators were significantly up-regulated and others were significantly down-regulated when comparing expression in PRAD and normal prostate patients. Next, we validated the protein expression level of the three up-regulated m⁶A-associated regulators (*AMDHD2*, *CKAP2L* and *RFC5*) between the normal prostate tissue and prostate cancer tissue from the HPA database, the PRAD patients showed moderate or weak staining signals, whereas it was not detected in all normal prostate tissue. These results provide a preliminary validation of the clinical results. However, due to the lack of large sample data in HPA database, further statistical analysis is needed to support the results. Nevertheless, there is no relevant literature report on the role of *AMDHD2* in cancer. Xiong *et al.* [33] have implied that *CKAP2L* expression is enhanced in lung adenocarcinoma tissues and is predictive of poor prognosis in lung adenocarcinoma patients, Wang *et al.* [34] have indicated that a high *RFC5* expression is associated with more aggressive malignant clinicopathological in lung cancer. However, the expression and mechanism of *CKAP2L* and *RFC5* in prostate cancer have not been reported, more *in vitro* and *in vivo* research is desirable to clarify the results.

Subsequently, a consensus clustering analysis on the PRAD patients was performed, and the PRAD patients could be classified into two subclasses. The results indicated significant differences in both clinicopathological characteristics and RFS. Next, we extracted its PPI association networks, and finally a network of nine nodes with one edge of protein interaction diagram were collected. These results mean that our current set of proteins may be rather small or essentially a random collection of proteins that are not very well connected. However, this does not necessarily mean that it is not a biologically meaningful selection of proteins; it could simply be that these proteins have not been studied very much and that their interactions might not yet be known to String database. Finally, the results of PCA also indicated significant differences between the two subclasses. All these results implied significant differences between the two subclasses.

Furthermore, GO functional and KEGG pathways enrichment analysis on the DEGs clusters between the two subclasses was utilized, respectively. In the GO analysis of the BP's domain for the up-regulated genes, the most cancer related annotations were found for ‘microtubule cytoskeleton organization’, at present, studies on disease and

disorder research in relation to the microtubule cytoskeleton organization in prostate cancer have been reported [35–37]. The next cancer related annotations was ‘chromosome segregation’, chromosomal instability is now known to be one of the hallmarks of cancer [38,39], Pan *et al.* [40] demonstrated that increasing chromosome segregation can lead to tumor cell growth in human prostate cancer. In the KEGG analysis for the up-regulated genes, the most cancer related pathways was ‘cell cycle’, which has been reported to be associated with malignancy, carcinogenesis and cellular transformation [41,42]. Most researches have demonstrated that cell cycle link to recurrence, growth and migration of PRAD [43–45]. The next cancer related signaling pathways was ‘p53 signaling pathway’, which is a typical cancer signaling pathway, Wan *et al.* [46] have demonstrated that PRAD cells exhibited low p53 expression, and the proliferation, migration and adhesion abilities of PRAD cells were promoted by inhibiting the activation of janus kinase (JNK) and extracellular signal-regulated kinase (ERK). More interestingly, enrichment analysis in the down-regulated genes revealed the involvement of ‘muscle contraction’ in both GO biological function and KEGG signaling pathway, as well as the involvement of ‘calcium signaling pathway’ and ‘cGMP-PKG signaling pathway’ in KEGG analysis. However, at present, there is a lack of research on this aspect in cancer. Therefore, we hypothesize that these genes might coordinate muscle contraction through these signaling pathways to regulate tumorigenesis. However, further *in vivo* and *in vitro* experiments are desirable to validate the downstream mechanism studies such as biological function and signaling pathways in PRAD.

Conclusion & future perspective

In summary, we comprehensively analyzed the expression pattern and prognostic model value of the m⁶A-associated regulators in PRAD patients, and nine m⁶A-associated variants related to the RFS of PRAD were ultimately obtained. Consensus clustering analysis on these regulators could divide PRAD into two subclasses with significant RFS prognosis, suggesting that these nine m⁶A-associated regulators could be used as biomarkers for clinical prognosis.

Summary points

- The first study to demonstrate the expression patterns and construction of a prognostic model of m⁶A-associated regulators in prostate adenocarcinoma (PRAD).
- A prognostic model with nine m⁶A-associated variants was ultimately obtained.
- All these nine m⁶A-associated regulators were related to the prognosis of PRAD, and could be used as biomarkers related to clinical prognosis.
- All these nine m⁶A-associated regulators were significantly regulated in PRAD.
- Six regulators were significantly down-regulated, and three regulators were significantly up-regulated in PRAD.
- These three up-regulated regulators were confirmed significant expression from the protein level in PRAD.
- PRAD patients could be classified into two risk groups with significant recurrence-free survival differences based on the prognostic model.
- PRAD patients could be divided into high- and low-risk subclasses with significant recurrence-free survival differences by consensus clustering analysis.

Author contributions

All authors contributed to study design and data acquisition, analysis or interpretation and drafting of this manuscript. All the authors had full access to all the data in the study and take responsibility for the integrity of the data and accuracy of data analysis.

Acknowledgments

The authors thanked the Department of Urology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology and Department of Urology, First Affiliated Hospital, School of Medicine, Shihezi University.

Financial & competing interests disclosure

This work was supported by the Key scientific and technological projects of Xinjiang production and Construction Corps (2018AB023). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations.

Data sharing statement

The data used to support the findings of this study are from the public database of The Cancer Genome Atlas (TCGA, cancer.genome.nih.gov), The Genotype-Tissue Expression (GTEx, www.gtexportal.org), m6AVar (m6avar.renlab.org) and Human Protein Atlas (HPA, www.proteinatlas.org).

References

Papers of special note have been highlighted as: • of interest

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68(6), 394–424 (2018).
- Nebbioso A, Tambaro FP, Dell'aversana C, Altucci L. Cancer epigenetics: moving forward. *PLoS Genet.* 14(6), e1007362 (2018).
- Pan Y, Ma P, Liu Y, Li W, Shu Y. Multiple functions of m(6)A RNA methylation in cancer. *J. Hematol. Oncol.* 11(1), 48 (2018).
- Wu Y, Sarkissyan M, Vadgama JV. Epigenetics in breast and prostate cancer. *Methods Mol. Biol.* 1238, 425–466 (2015).
- Zhao L, Duan YT, Lu P *et al.* Epigenetic targets and their inhibitors in cancer therapy. *Curr. Top. Med. Chem.* 18(28), 2395–2419 (2018).
- Morena F, Argentati C, Bazzucchi M, Emiliani C, Martino S. Above the epitranscriptome: RNA modifications and stem cell identity. *Genes (Basel)* 9(7), 329 (2018).
- Dai D, Wang H, Zhu L, Jin H, Wang X. N⁶-methyladenosine links RNA metabolism to cancer progression. *Cell Death Dis.* 9(2), 124 (2018).
- Batista PJ. The RNA modification N⁶-methyladenosine and its implications in human disease. *Genomics Proteomics Bioinformatics* 15(3), 154–163 (2017).
- Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149(7), 1635–1646 (2012).
- Chai RC, Wu F, Wang QX *et al.* m⁶A RNA methylation regulators contribute to malignant progression and have clinical prognostic impact in gliomas. *Aging (Albany NY)* 11(4), 1204–1225 (2019).
- **Evaluates the relationship between m⁶A RNA methylation regulators and malignant progression clinical prognostic impact in gliomas.**
- Lin S, Choe J, Du P, Triboulet R, Gregory RI. The m⁶A methyltransferase METTL3 promotes translation in human cancer cells. *Mol. Cell* 62(3), 335–345 (2016).
- Zhang C, Samanta D, Lu H *et al.* Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m⁶A-demethylation of NANOG mRNA. *Proc. Natl Acad. Sci USA* 113(14), E2047–2056 (2016).
- Mounir M, Lucchetta M, Silva TC *et al.* New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. *PLoS Comput. Biol.* 15(3), e1006701 (2019).
- **This package can able to access The National Cancer Institute (NCI) Genomic Data Commons (GDC) thorough its GDC Application Programming Interface (API) to search, download and prepare relevant data for analysis in R, including the 'The Cancer Genome Atlas' and 'Genotype-Tissue Expression' database.**
- Zheng Y, Nie P, Peng D *et al.* m⁶AVar: a database of functional variants involved in m⁶A modification. *Nucleic Acids Res.* 46(D1), D139–D145 (2018).
- **The m⁶AVar database is specifically designed to collect functional variants that can significantly alter the status of m⁶A, including the m⁶A-associated cancer somatic mutations.**
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33(1), 1–22 (2010).
- **This package fits least absolute shrinkage and selection operator and elastic-net model paths for regression, logistic and multinomial regression using coordinate descent.**
- Uhlen M, Fagerberg L, Hallstrom BM *et al.* Proteomics. tissue-based map of the human proteome. *Science* 347(6220), 1260419 (2015).
- Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26(12), 1572–1573 (2010).
- **This package provides an algorithm for determining cluster count and membership by stability evidence in unsupervised analysis.**
- Szklarczyk D, Gable AL, Lyon D *et al.* STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47(D1), D607–D613 (2019).
- **This package provides an algorithm for determining cluster count and membership by stability evidence in unsupervised analysis.**

19. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1), 139–140 (2010).
20. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16(5), 284–287 (2012).
- **This R package implements methods to analyze and visualize functional profiles of gene and gene clusters.**
21. Roundtree IA, Evans ME, Pan T, He C. Dynamic RNA modifications in gene expression regulation. *Cell* 169(7), 1187–1200 (2017).
22. Deng X, Su R, Feng X, Wei M, Chen J. Role of N⁶-methyladenosine modification in cancer. *Curr. Opin. Genet. Dev.* 48, 1–7 (2018).
23. Meyer KD, Jaffrey SR. The dynamic epitranscriptome: n⁶-methyladenosine and gene expression control. *Nat. Rev. Mol. Cell Biol.* 15(5), 313–326 (2014).
24. Alarcón CR, Goodarzi H, Lee H, Liu X, Tavazoie S, Tavazoie SF. HNRNPA2B1 is a mediator of m⁶A-dependent nuclear RNA processing events. *Cell* 162(6), 1299–1308 (2015).
25. Chen T, Hao Y-J, Zhang Y *et al.* m⁶A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. *Cell Stem Cell* 16(3), 289–301 (2015).
26. Huarte M. The emerging role of lncRNAs in cancer. *Nat. Med.* 21(11), 1253–1261 (2015).
27. Hayes J, Peruzzi PP, Lawler S. MicroRNAs in cancer: biomarkers, functions and therapy. *Trends Mol. Med.* 20(8), 460–469 (2014).
28. Meng Y, Li S, Gu D *et al.* Genetic variants in m⁶A modification genes are associated with colorectal cancer risk. *Carcinogenesis* 41(1), 8–17 (2020).
- **Evaluates the relationship between genetic variants in m⁶A modification and predictors risk in colorectal cancer.**
29. Li E, Wei B, Wang X, Kang R. METTL3 enhances cell adhesion through stabilizing integrin β1 mRNA via an m⁶A-HuR-dependent mechanism in prostatic carcinoma. *Am. J. Cancer Res.* 10(3), 1012–1025 (2020).
30. Yuan Y, Du Y, Wang L, Liu X. The m⁶A methyltransferase METTL3 promotes the development and progression of prostate carcinoma via mediating MYC methylation. *J. Cancer* 11(12), 3588–3595 (2020).
31. Ma X-X, Cao Z-G, Zhao S-L. m⁶A methyltransferase METTL3 promotes the progression of prostate cancer via m⁶A-modified LEF1. *Eur. Rev. Med. Pharmacol. Sci.* 24(7), 3565–3571 (2020).
32. Barros-Silva D, Lobo J, Guimarães-Teixeira C *et al.* VIRMA-dependent N⁶-methyladenosine modifications regulate the expression of long non-coding RNAs CCAT1 and CCAT2 in prostate cancer. *Cancers (Basel)* 12(4), 771 (2020).
33. Xiong G, Li L, Chen X *et al.* Up-regulation of CKAP2L expression promotes lung adenocarcinoma invasion and is associated with poor prognosis. *Onco Targets Ther.* 12, 1171–1180 (2019).
34. Wang M, Xie T, Wu Y *et al.* Identification of RFC5 as a novel potential prognostic biomarker in lung cancer through bioinformatics analysis. *Oncol. Lett.* 16(4), 4201–4210 (2018).
35. Levrier C, Rockstroh A, Gabrielli B *et al.* Discovery of thalictuberine as a novel antimetabolic agent from nature that disrupts microtubule dynamics and induces apoptosis in prostate cancer cells. *Cell Cycle* 17(5), 652–668 (2018).
36. Mukhtar E, Adhami VM, Sechi M, Mukhtar H. Dietary flavonoid fisetin binds to beta-tubulin and disrupts microtubule dynamics in prostate cancer cells. *Cancer Lett.* 367(2), 173–183 (2015).
37. Ho SM, Rao R, To S, Schoch E, Tarapore P. Bisphenol A and its analogues disrupt centrosome cycle and microtubule dynamics in prostate cancer. *Endocr. Relat. Cancer* 24(2), 83–96 (2017).
38. Bochtler T, Kartal-Kaess M, Granzow M *et al.* Micronucleus formation in human cancer cells is biased by chromosome size. *Genes Chromosomes Cancer* 58(6), 392–395 (2019).
39. Zhang BN, Bueno Venegas A, Hickson ID, Chu WK. DNA replication stress and its impact on chromosome segregation and tumorigenesis. *Semin. Cancer Biol.* 55, 61–69 (2019).
40. Pan HW, Su HH, Hsu CW, Huang GJ, Wu TT. Targeted TPX2 increases chromosome missegregation and suppresses tumor cell growth in human prostate cancer. *Onco Targets Ther.* 10, 3531–3543 (2017).
41. Wenzel ES, Singh ATK. Cell-cycle checkpoints and aneuploidy on the path to cancer. *In Vivo* 32(1), 1–5 (2018).
42. Jacobberger JW, Sramkoski RM, Stefan T, Woost PG. Multiparameter cell cycle analysis. *Methods Mol. Biol.* 1678, 203–247 (2018).
43. Nam RK, Benatar T, Wallis CJD *et al.* MicroRNA-139 is a predictor of prostate cancer recurrence and inhibits growth and migration of prostate cancer cells through cell cycle arrest and targeting IGF1R and AXL. *Prostate* 79(12), 1422–1438 (2019).
44. Georgescu C, Corbin JM, Thibivilliers S *et al.* A TMEFF2-regulated cell cycle derived gene signature is prognostic of recurrence risk in prostate cancer. *BMC Cancer* 19(1), 423 (2019).
45. Rubicz R, Zhao S, April C *et al.* Expression of cell cycle-regulated genes and prostate cancer prognosis in a population-based cohort. *Prostate* 75(13), 1354–1362 (2015).
46. Wan J, Zhang J, Zhang J. Expression of p53 and its mechanism in prostate cancer. *Oncol. Lett.* 16(1), 378–382 (2018).

