# Identification of core genes associated with prostate cancer progression and outcome via bioinformatics analysis in multiple databases

Yutao Wang[1],[*], Jianfeng Wang[1],[*], Kexin Yan[2], Jiaxing Lin[1], Zhenhua Zheng[1] and Jianbin Bi[1]

[1] Department of Urology, The First Hospital of China Medical University, Shenyang, China
[2] Department of Dermatology, The First Hospital of China Medical University, Shenyang, China
[*] These authors contributed equally to this work.

**Abstract:** The morbidity and mortality of prostate carcinoma has increased in recent years and has become the second most common ale malignant carcinoma worldwide. The interaction mechanisms between different genes and signaling pathways, however, are still unclear.
**Methods:** Variation analysis of GSE38241, GSE69223, GSE46602 and GSE104749 were realized by GEO2R in Gene Expression Omnibus database. Function enrichment was analyzed by DAVID.6.8. Furthermore, the PPI network and the significant module were analyzed by Cytoscape, STRING and MCODE.GO. Pathway analysis showed that the 20 candidate genes were closely related to mitosis, cell division, cell cycle phases and the p53 signaling pathway. A total of six independent prognostic factors were identified in GSE21032 and TCGA PRAD. Oncomine database and The Human Protein Atlas were applied to explicit that six core genes were over expression in prostate cancer compared to normal prostate tissue in the process of transcriptional and translational. Finally, gene set enrichment were performed to identified the related pathway of core genes involved in prostate cancer.
**Result:** Hierarchical clustering analysis revealed that these 20 core genes were mostly related to carcinogenesis and development. CKS2, TK1, MKI67, TOP2A, CCNB1 and RRM2 directly related to the recurrence and prognosis of prostate cancer. This result was verified by TCGA database and GSE21032.
**Conclusion:** These core genes play a crucial role in tumor carcinogenesis, development, recurrence, metastasis and progression. Identifying these genes could help us to understand the molecular mechanisms and provide potential biomarkers for the diagnosis and treatment of prostate cancer.

**Subjects** Bioinformatics, Genomics, Urology
**Keywords** Prostate cancer, Prognosis factor, Biomarkers, GEO, TCGA

## INTRODUCTION

Prostate cancer is the second most common male malignancy tumor in the world (*Farhood et al., 2019*). The morbidity and mortality of prostate cancer has surpassed bladder cancer and kidney cancer. It is now the most common tumor in the adult urology

department in China (*Mayer et al., 2001*). There are several factors can increase the risk of prostate cancer (for example, family factors, bald, gonorrhea, smoking; *Lian et al., 2015*; *Rao et al., 2015*; *Islami et al., 2014*), and genetic predisposition is considered to be one of the important factors of the incidence of prostate cancer (*Ju-Kun et al., 2016*; *Jansson et al., 2012*; *Hemminki, 2012*). At present, 100 susceptibility loci were identified, the mechanism of prostate cancer is incomprehensible, although a large number of studies have been conducted on the development and recurrence of prostate cancer. Prostate-specific antigen (PSA) test has been used to assist the diagnosis of prostate cancer. Early PSA detection, however, could lead to over diagnosis and over treatment of prostate cancer. More recently, advances in high-throughput sequencing and screening techniques have enabled us to screen differently expressed genes at the same time. Therefore, this article aimed to identify the core protein coding genes related to the progression of cancer recurrence, metastasis and prognosis by bioinformatics analysis. As errors in an individual dataset are unavoidable, multiple datasets were analyzed. False positive results could be ignored by independent mRNA microarray analysis; false negative results have been ignored by multiple mRNA microarrays intersection analysis. Thus, four kinds of combinations were established between four datasets downloaded from Gene Expression Omnibus (GEO), differentially expressed genes (DEGs) were identified between cancer and non cancer tissues when the screening criteria met any three of the four datasets (*Li et al., 2017*). Afterwards we obtained six core differential expression genes which were verified in TCGA PRAD, GSE21032 and other datasets, besides these six core genes acted as independent prognostic factors and positively correlated with Gleason score (*Katarzyna, Patrycja & Maciej, 2015*). The six real core genes were thereby identified as potential biomarkers for prostate cancer.

# MATERIALS AND METHODS

## Microarray data

Five prostate cancer microarray datasets were obtained from NCBI GEO (https://www.ncbi.nlm.nih.gov/geo/) (*Edgar, Domrachev & Lash, 2002*): GSE104749 (*Shan et al., 2017*), GSE38241 (*Aryee et al., 2013*), GSE69223 (*Meller et al., 2016*) and GSE46602 (*Mortensen et al., 2015*), GSE21032 (*Taylor et al., 2010*). The platform for GSE104749 was GPL570 Affymetrix Human Genome U133 Plus 2.0 Array which contained 4 PCa excision tissue and 4 noncancerous excision tissue samples. The platform for GSE38241 was GPL4133 Agilent-014850 Whole Human Genome microarray 4 × 44 K G4112F (Feature Number version) which contained 21 PCa excision tissue and 18 noncancerous excision tissue samples. The platform for GSE69223 was GPL 570 Affymetrix Human Genome U133 Plus 2.0 Array which contained 15 PCa excision tissue samples and 15 noncancerous excision tissue samples. The platform for GSE46602 was GPL570 Affymetrix Human Genome U133 Plus 2.0 Array which contained 36 excision tissue and 14 noncancerous tissue samples. The platform for GSE21032 was GPL5188 which contained 150 prostate excision tissue samples and 29 normal adjacent benign prostate excision tissue samples. Meanwhile we also downloaded clinical information and gene matrix from TCGA

database (https://genome-cancer.ucsc.edu/) (*Cancer Genome Atlas Research Network TCGA, 2015*) which contained 52 normal prostate tissue samples and 495 prostate cancer samples.

### Identification of DEGs

The GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r) was used to screen DEGs between prostate carcinoma samples and noncancerous samples GSE104749, GSE38241, GSE69223, GSE46602 datasets. GEO2R is a tool which allows users to obtain DEGs by comparing different groups. The DEGs are screened and sorted by significance. The genes with |log2FC| (fold change) ≥ 1 and *P*-value < 0.01 were considered to be DEGs.

### GO and pathway enrichment analysis

The Database for Annotation, Visualization and Integrated Discovery (DAVID, v6.8) is a function enrichment tool which supplies biological explanations of gene lists and proteomic studies from high-throughput sequencing (*Huang et al., 2007*). Higher-order functions of cells and organisms were derived from KEGG (http://www.genome.ad.jp/kegg/) databases (*Kanehisa, 2002*). The biological process, molecular function and cellular component were analyzed in GO (http://www.geneontology.org) (*Ashburner et al., 2000*). A *P* value of < 0.05 was the criterion for significance.

### Conduction of protein–protein interaction network

The direct and indirect interactions between proteins were established by STRING database (http://string-db.org/) (*Franceschini et al., 2013*). The PPI network was analyzed by STRING database with a criterion of combined score >0.4 considered to be a significant result. The biological network visualization of the protein interactions was revealed by the open source software Cytoscape 3.7.1 (*Smoot et al., 2011*). The MCODE in Cytoscape was used to the 20 core genes and significant module (*Bandettini et al., 2012*). Afterwards the PPI network and the significant module were established with the criteria (MCODE scores > 5, degree cutoff = 2, node score cutoff = 0.2, Max depth = 100 and *k*-score = 2).

### Hub genes selection and prognosis analysis

The candidate core genes were identified with degrees ≥20 of the most significant model in PPI network. The co-expression network of the candidate genes was established by the online platform cBioPortal (http://www.cbioportal.org) (*Gao et al., 2013*). The phenotype analysis of the core genes was performed by a heat map that was conducted by UCSC (http://genome-cancer.ucsc.edu) (*Haeussler et al., 2019*). Afterwards, to further screen out the independent prognosis factors of prostate cancer, GSE21032 log2 mRNA expression data and the clinical information were download from cBioPortal (http://cbio.mskcc.org/cancergenomics/prostate/data) and (https://www.cbioportal.org/study/clinicalData?id=prad_mskcc). We analyzed the correlation between candidate key genes and phenotype in the GSE21032 data performed by box plots, KM curve and ROC curve, *P* value < 0.05 was the criterion for significance. In this way, we identified the real core genes and verified their important value in the large sample TCGA.

## External data set evaluation and verification

The core protein coding genes were identified, the prognosis and clinical value were illustrated in the above study. Furthermore, the interaction between the core genes and metastasis state were analyzed by Oncomine (http://www.oncomine.com) (*Vanaja et al., 2003*; *Grasso et al., 2012*; *Varambally et al., 2005*). The Human Protein Atlas (HPA) (http://www.proteinatlas.org/) is an open source database to explore the human proteome. HPA combines various omics techniques to map all human proteins in cells, tissues and organs and the expression of core genes were evaluated on transcriptional and translational level. The analysis of variance was applied to show the correlativity between the expression of core genes and clinical stage to reveal the expression difference among different stages, the clinical data was from TCGA clinical information.

## Gene set enrichment analysis

In set TCGA, samples were divided into two groups based on the expression of core genes. To predict the function and effect of core genes, gene set enrichment (GSEA) (http://software.broadinstitute.org/gsea/index.jsp) (*Subramanian et al., 2005*) was applied to identified the gene set that expressed relatively with our core genes, and analyzed the pathway the gene set involved in. Pathway with $P < 0.05$ was considered to be significance. Afterwards, we identified the most significance pathway by taking the intersection of core genes. The results were performed by "ggplot2" packages (*Ito & Murphy, 2013*) in R 3.6.2.
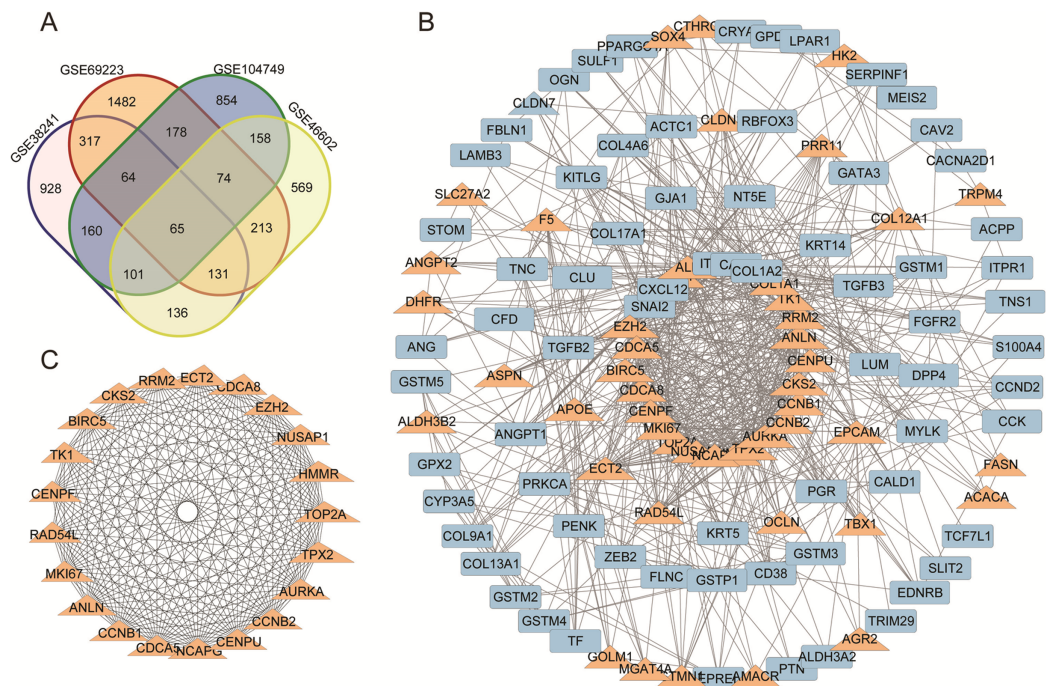
# RESULT

## Identification of DEGs in PCa

A total of 298 downregulated genes and 137 upregulated DEGs were screened, as shown in the Venn diagram (Fig. 1A). Expression datasets GSE46602, GSE104749, GSE38241 and GSE69223 were obtained from the GEO database. GEO2R was used to screen the DEGs with the criteria of |log2FC| (fold change) $\geq 1$ and $P$-value $< 0.01$. Afterward, we got 1,953 DEGs in GSE46602, 2,513 DEGs in GSE104749, 2,682 DEGs in GSE38241 and 2,639 DEGs in GSE69223. Considering the errors between the four datasets, we used the genes that met the screening criteria of any three datasets for the next analysis.

TCGA prostate cancer gene matrix was transformed by log2(exp + 1).

## KEGG and GO enrichment analyses of DEGs

DAVID was used to analyze the function classification of the 298 downregulated genes and 137 upregulated DEGs obtained by intersection. Gene Ontology results showed DEGs were significantly enriched in cell and biological adhesion, mesenchymal cell differentiation and the development of biological process. For the cellular component groups, the extracellular region and the matrix were significantly enriched. In the group analyzed for molecular function, the glutathione transferase activity, extracellular matrix structural constituent identical protein binding, and general identical protein binding were significantly enriched. KEGG pathway analysis revealed that the DEGs were mainly enriched in focal adhesion and in the glutathione metabolism pathway (Table 1).

Figure 1 (A) DEGs Venn diagram of four prostate cancer microarray data. (B) PPI network for DEGs. (C) The most significant module of DEGs. (A) DEGs Venn diagram of GSE38241, GSE69223, GSE104749, GSE46602. As a result, 298 downregulated genes and 137 upregulated DEGs were screened, as shown in the Venn diagram. (B) The protein and protein interactions (PPI) 340 DEGs in 435 DEGs were performed in protein network and 167 DEGs were shown with degree >7. The upregulated genes were marked with orange,and the downregulated gene were marked with blue. (C) The most statistically significant module was acquired using Cytoscape.

Full-size ⊡ DOI: 10.7717/peerj.8786/fig-1

## PPI and module analysis

We acquired the PPI network and identified the most significant model based on the 298 downregulated genes and 137 upregulated DEGs. Cytoscape software and the online database STRING (available online: https://string-db.org/) were used to screen core genes. 340 DEGs in 435 DEGs were performed in protein network and 167 DEGs were shown with degree >7 (Fig. 1B). The most statistically significant module was acquired using Cytoscape (Fig. 1C). The KEGG and GO enrichment of this module were analyzed using DAVID. Results showed that the genes in the most significant module were mainly enriched in the functions of cell cycle phases, mitosis, cell division, the microtubule cytoskeleton, the p53 signaling pathway and Oocyte meiosis (Table 2).

## Candidate gene selection, analysis

A total of 20 core genes with a degree ≥20 selected by MCODE were obtained from the protein–protein network, and they were considered to be candidate core genes. The co-expression network of core genes was analyzed by cBioPortal (Fig. 2A). Heatmap showed that the core genes in cancer groups upregulated more significantly than in normal groups by UCSC (https://xena.ucsc.edu) (Fig. 2C). In addition, the core genes were

**Table 1 GO and KEGG pathway enrichment analysis of DEGs in prostate cancer samples.** KEGG pathway analysis revealed that the DEGs were mainly enriched in focal adhesion and Glutathione metabolism pathway.

| Pathway ID | Term | Count | P-value |
|---|---|---|---|
| GO:0007155 | Cell adhesion | 39 | 1.86E−06 |
| GO:0007155 | Biological adhesion | 39 | 1.88E−06 |
| GO:0048762 | Mesenchymal cell differentiation | 9 | 2.52E−05 |
| GO:0014031 | Mesenchymal cell development | 9 | 2.52E−05 |
| GO:0044421 | Extracellular region part | 63 | 3.08E−14 |
| GO:0031012 | Extracellular matrix | 32 | 6.05E−11 |
| GO:0005578 | Proteinaceous extracellular matrix | 29 | 9.94E−10 |
| GO:0005578 | Extracellular region | 87 | 1.37E−09 |
| GO:0004364 | Glutathione transferase activity | 6 | 7.31E−05 |
| GO:0005201 | Extracellular matrix structural constituent | 10 | 1.57E−04 |
| hsa04510 | Focal adhesion | 17 | 4.85E−05 |
| hsa04510 | Drug metabolism | 9 | 1.64E−04 |
| hsa04510 | Glutathione metabolism | 8 | 2.54E−04 |

Note:
GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, differentially expressed genes.

**Table 2 GO and KEGG pathway enrichment analysis of DEGs in the most significant module.** The genes in the most significant module were mainly enriched in the function of cell cycle phase, mitosis, cell division, microtubule cytoskeleton, spindle, p53 signaling pathway, Oocyte meiosis.

| Pathway ID | Term | Count | P-value |
|---|---|---|---|
| GO:0000279 | M phase | 14 | 6.27E−18 |
| GO:0022403 | Cell cycle phase | 14 | 1.27E−16 |
| GO:0007067 | Mitosis | 11 | 4.12E−14 |
| GO:0000280 | Nuclear division | 11 | 4.12E−14 |
| GO:0051301 | Cell division | 10 | 4.04E−11 |
| GO:0015630 | Microtubule cytoskeleton | 9 | 1.90E−07 |
| GO:0005819 | Spindle | 6 | 1.04E−06 |
| GO:0051276 | Chromosome organization | 8 | 1.65E−06 |
| GO:0044430 | Cytoskeletal part | 9 | 1.22E−05 |
| GO:0003682 | Chromatin binding | 4 | 6.21E−04 |
| hsa04115 | p53 Signaling pathway | 3 | 0.003543755 |
| hsa04114 | Oocyte meiosis | 3 | 0.009071242 |

Note:
GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, differentially expressed genes.

considered to have a close relation to the Gleason score (Fig. 2B). We adjusted the highest color according to 100% saturation parameters of log2 (norm_count + 1) ≥ 10.4 and the lowest color according to 100% saturation parameters of log2 (norm_count + 1) ≤ 2.65 (Fig. 2B). We compared solid normal tissue to primary tumor tissue (Figs. 2B and 2C). We found that CCNB1, TPX2, CENPF, TOP2A, MKI67, ECT2, TK1, RRM2, NUSAP1,
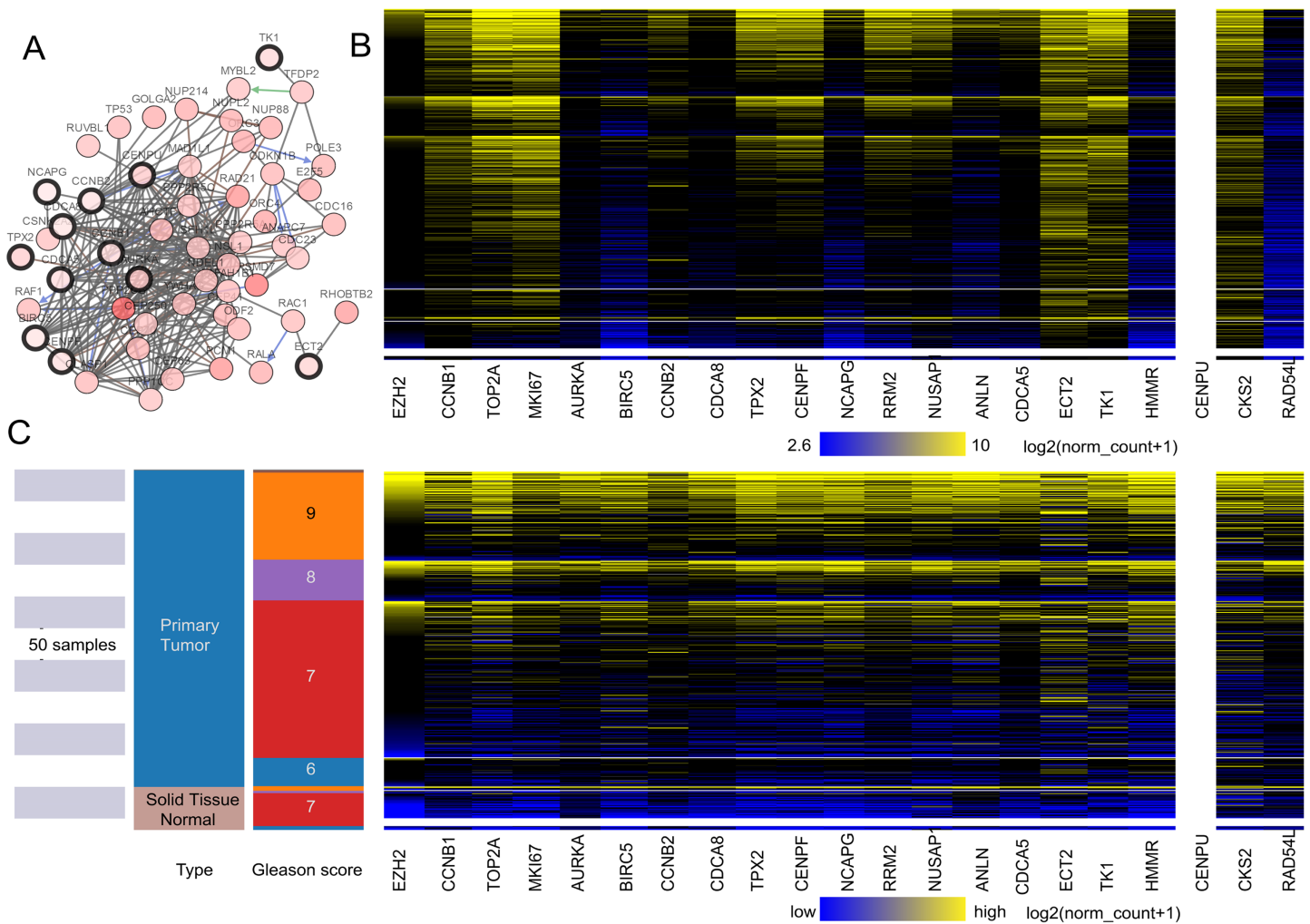
**Figure 2** (A) The interaction network between the hub genes and their co-expression genes. (B and C) Heat map of hub genes by UCSC. (A) The interaction network between the hub genes and their co-expression genes by cBioPortal. Nodes with a black outline represent hub genes. Nodes with thin black genes represent co-expression genes. (B) The color parameters of C were adjusted as max color 100% saturation of log2 (norm_count + 1) = 10.4; min color 100% saturation of log2(norm_count + 1) = 2.65. (C) Heat map of hub genes by UCSC. The samples of pink are normal tissue. The cancer samples are blue. High expressed genes were marked yellow. The low expressed genes were blue.

Full-size ☑ DOI: 10.7717/peerj.8786/fig-2

CKS2 were overexpressed consistently in the 568 TCGA Prostate Cancer (PRAD) samples. Therefore, these genes were closely related to carcinogenesis and stage.

## Selection real core genes

To further explore the relationship to phenotype and prognosis status, we performed differential analysis, disease-free survival analysis and ROC curve analysis in GSE21032 (Figs. 3A–3R). CCNB1, CKS2, RRM2, MKI67, TK1 and TOP2A were considered to be core genes with the lowest $P$ value in the above 20 candidate genes. The results showed that MKI67 had the best prognostic value and the most significant $P$ value ($P = 0.0024$). In addition, CCNB1, CKS2, RRM2, TK1 and TOP2A also showed closely clinical correlation. Based on the above analysis, the six genes were predicted to be the core factors
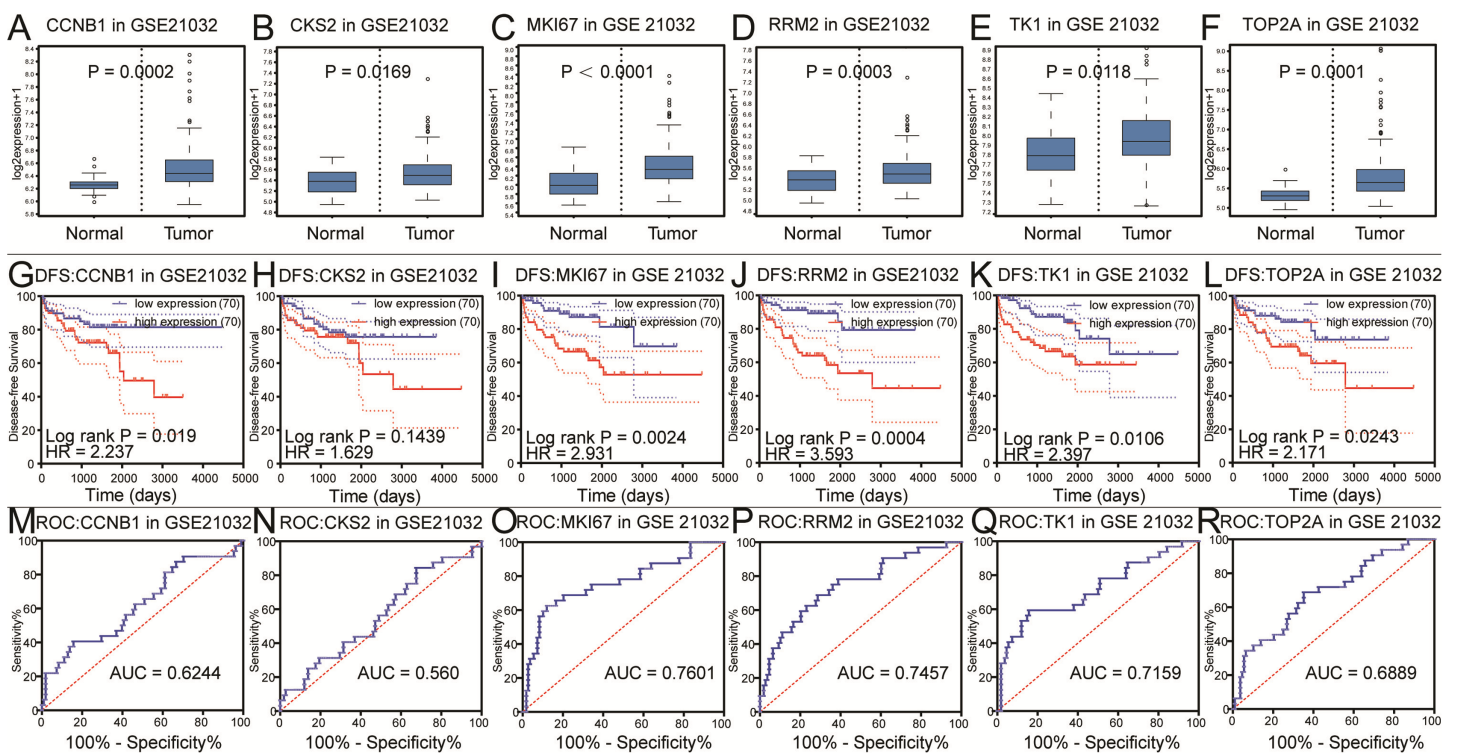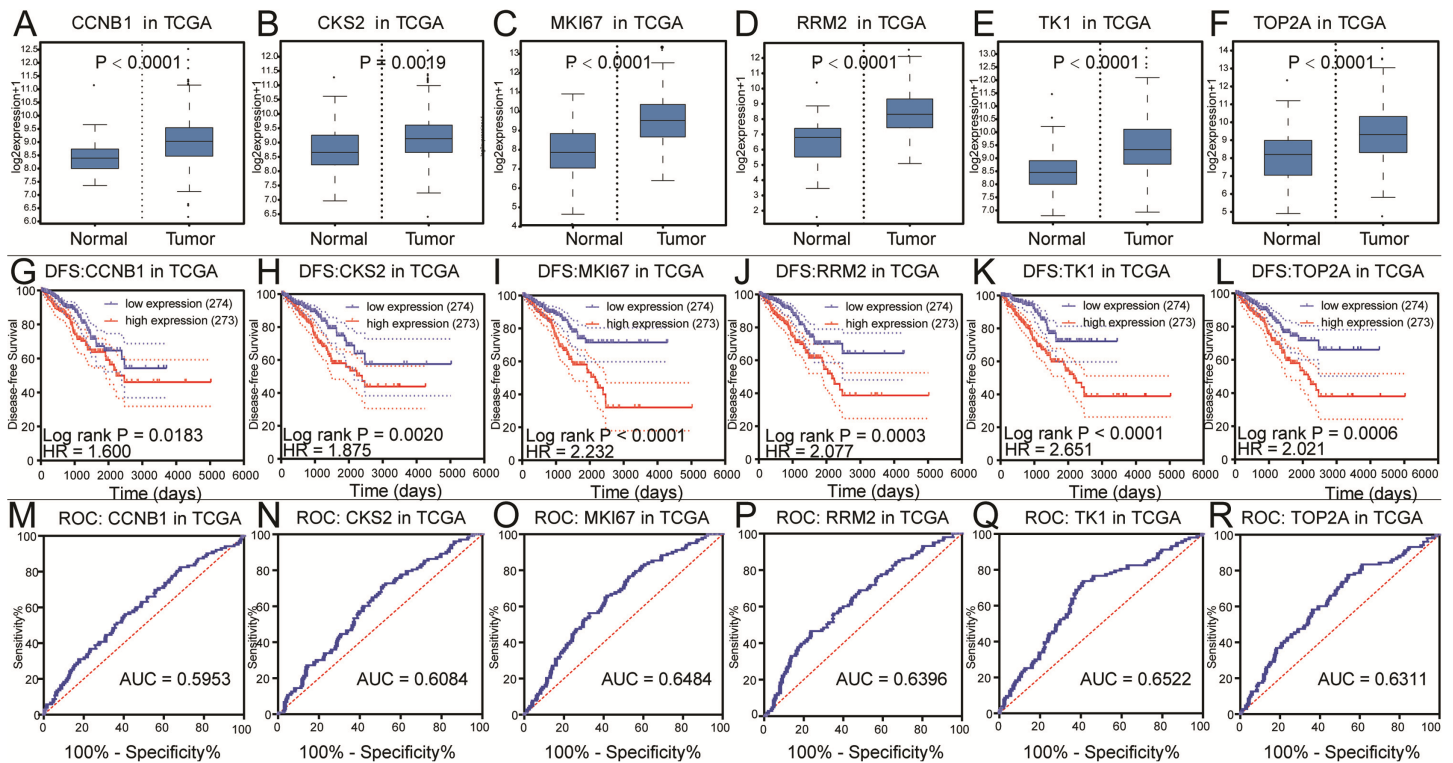
**Figure 3 Differential analysis, disease-free survival analysis and ROC curve analysis in GSE21032.** (A–F) The box plots of CCNB1, CKS2, MKI67, RRM2, TK1, TOP2A from GSE21032. The six genes were found to be risk factors in GSE21032. (G–L) The disease-free survival KM curve of the six core genes were analysis by GSE21032, which showed prognosis value of prostate cancer patients. The Log-rank *P* value < 0.05 was considered to be significance. We recommended CKS2 as prognosis factors because the survival difference in 5 years is significance. (M–R) The ROC curve and AUC of the six core genes were showed. Full-size 🖼 DOI: 10.7717/peerj.8786/fig-3

affecting prostate cancer. To determine our prediction, we found CCNB1, CKS2, MKI67, RRM2, TK1 and TOP2A acted as independent prognostic factors in TCGA prostate cancer (Figs. 4A–4R). In order to verify the differential expression of key factors at the transcriptional and translational levels, we found cases of immunohistochemistry of core factors in the HPA database. In contrast to normal prostate tissue, we found that the cancer group stained deeper. At the same time, the expression difference of core genes was found between local lesions and metastatic patients in different external data sets of Oncomine database (Figs. 5A–5L). These findings clarified that core factors played an important role in the entire process of prostate cancer.

## Clinical relevance and GSEA

CCNB1, CKS2, MKI67, RRM2, TK1 and TOP2A were found with positive clinical correlation to GLEASON SCORE (Figs. 6A–6F), and we could clearly identify the core factors as cancer risk factors, the expression level increased as Gleason score elevated. GSEA analysis suggested that cell cycle, DNA replication, GNRH signaling pathway, P53_signaling_pathway, were enriched in CKS2, MKI67, RRM2, TK1 and TOP2A high expression group jointly (Figs. 7A–7E). The *P*-value of core genes were listed (Table 3). Therefore, the core genes were positively related to the four cancer related pathways jointly.

**Figure 4 Differential analysis, disease-free survival analysis and ROC curve analysis in TCGA.** (A–F) The box plots of CCNB1, CKS2, MKI67, RRM2, TK1, TOP2A from TCGA. The six core genes were identified to be risk factors in TCGA. (G–L) The disease-free survival KM curve of the six core genes were analysis by TCGA, which demonstrated prognosis value of prostate cancer patients. The Log-rank *P* value < 0.05 was considered to be significance. (M–R) The ROC curve and AUC of the six core genes were showed. Full-size ◩ DOI: 10.7717/peerj.8786/fig-4
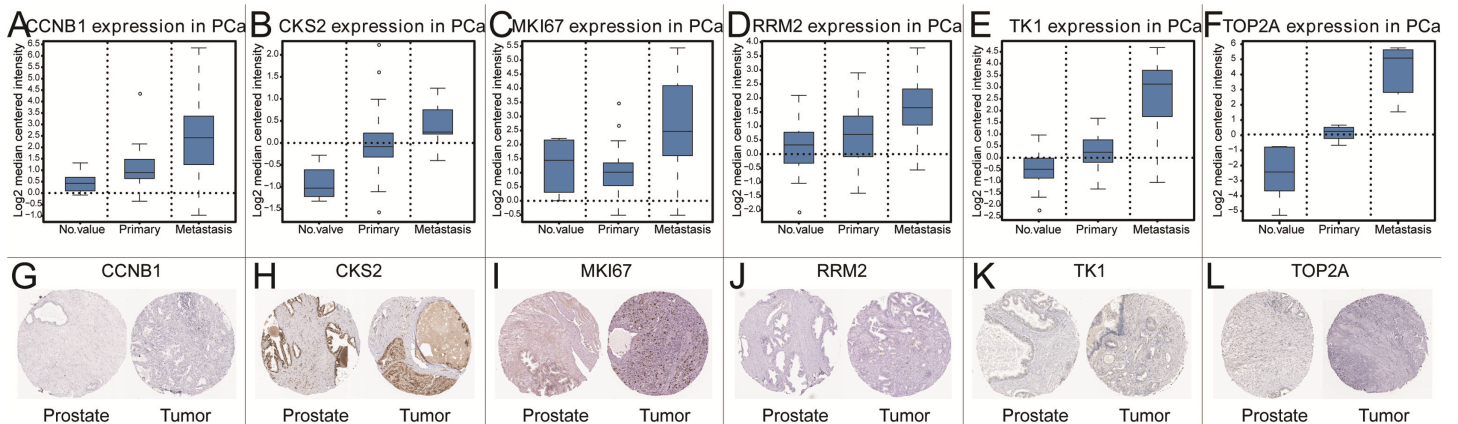


**Figure 5 The gene expression of CCNB1, CKS2, MKI67, RRM2, TK1, TOP2A on transcriptional and translational level.** Verification of the expression of core genes on transcriptional and translational level by Oncomine database and The Human Protein Atlas database (immunohistochemistry) CCNB1 (A and G), CKS2 (B and H), MKI67 (C and I), RRM2 (D and J), TK1 (E and K), TOP2A (F and L). Full-size ◩ DOI: 10.7717/peerj.8786/fig-5

# DISCUSSION

Despite the advances that have been made in understanding the molecular processes at the onset and in the progression of this disease, PCa still remains high morbidity and
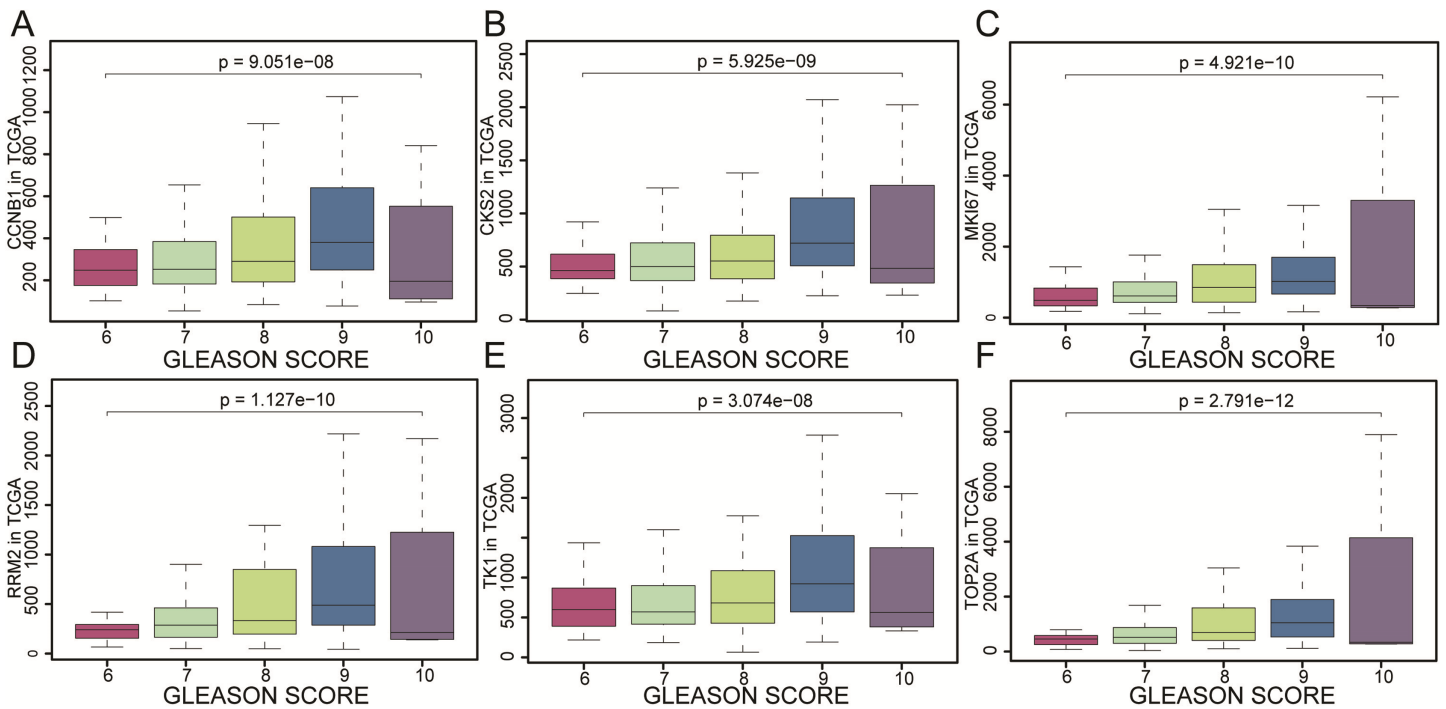
**Figure 6 The gene expression of CCNB1, CKS2, MKI67, RRM2, TK1, TOP2A as represented by different Gleason scores.** The gene expression of (A) CCNB1, (B) CKS2, (C) MKI67, (D) RRM2, (E) TK1, (F) TOP2A as represented by different Gleason scores. The results were evaluated by variance analysis. Full-size ◪ DOI: 10.7717/peerj.8786/fig-6
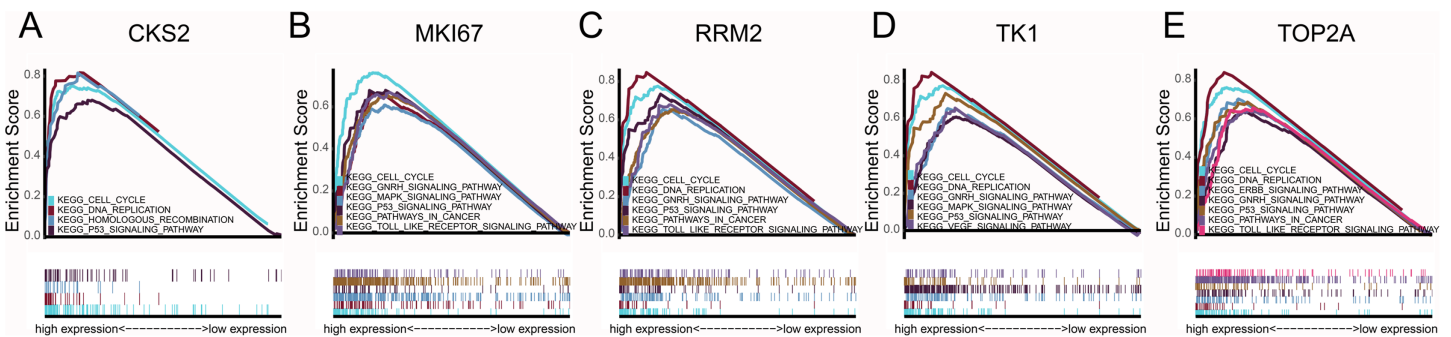


**Figure 7 GSEA pathway enrichment analysis.** GSEA pathway enrichment analysis of high expression groups of (A) CKS2, (B) MKI67, (C) RRM2, (D) TK1, (E) TOP2A. Pathway with FDR < 0.05 were considered to be significance. Full-size ◪ DOI: 10.7717/peerj.8786/fig-7

mortality, particularly in low-income countries (*Hernández et al., 2019*). However, the formation mechanisms of PCa are incomprehensive. Therefore, we excavated potential biomarkers for diagnosis and treatment (*Yang et al., 2018*).

In the present study, 298 downregulated and 137 upregulated genes mainly enriched in cell adhesion, biological adhesion, regulation of cell proliferation and oxidation reduction. Epithelial cell adhesion molecule (Ep-CAM) is considered to have a critical role in carcinogenesis and cell proliferation (*Tai et al., 2007*). In addition, oxidoreductase activity often plays a crucial role in antioxidant defense and encodes tumor suppressors that are active in tumorigenesis (*Li et al., 2017*). Moreover, in the prostate carcinoma PC-3 cell

**Table 3 GSEA results of core genes in TCGA PRAD.** The *P*-value of the P53_signaling_pathway, GNRH signaling pathway, DNA replication, cell cycle in the 5 core genes were listed.

| | P53_signaling_pathway | | GNRH_signaling_pathway | | DNA_replication | | Cell_cycle | |
|---|---|---|---|---|---|---|---|---|
| | NOM *P*-val | FDR *q*-val | NOM *P*-val | FDR *q*-val | NOM *P*-val | FDR *q*-val | NOM *P*-val | FDR *q*-val |
| RRM2 | <0.001 | 0.0012 | <0.001 | 4.65E−04 | <0.001 | 0.001 | <0.001 | 6.04E−04 |
| CKS2 | <0.001 | 0.0054 | 0.125 | 0.1249 | <0.001 | 0.0059 | <0.001 | 0.00254688 |
| MKI67 | <0.001 | 0.0026 | <0.001 | 9.34E−04 | <0.001 | 0.0023 | <0.001 | 0.001120284 |
| TK1 | <0.001 | 0.0124 | 0.0653 | 0.0708 | <0.001 | 0.0049 | <0.001 | 0.003169282 |
| TOP2A | 0.001992032 | 0.0045 | <0.001 | 0.0029 | <0.001 | 0.004227 | <0.001 | 0.00279397 |

model, the action of the gastrin releasing peptide (GRP) analog, bombesin (BN), on the activation of focal adhesion kinase (FAK) and its invasiveness suggests that this kinase might favor metastasis (*Lacoste, Aprikian & Chevalier, 2005*). In conclusion, these biological functions are closely related to the development and progression of prostate cancer.

Afterwards, a total of 20 candidate core genes were screened with degree ≥20. The full name and the function of the 20 core genes are listed (Table 4). CCNB1, CKS2, MKI67, RRM2, TK1, TOP2A were identified as real core genes by GSE21032 and validated in TCGA. The overexpression of CKS2 corresponds to metastasis and prognosis in various malignancies such as breast cancer, liver cancer and PCa (*Yu, Zhong & Qiao, 2013*). Other research showed that the forced expression of CKS2 has a positive correlation to cell growth, and also protects the cells from apoptosis (*Lan et al., 2008*). Thymidine kinase 1 (TK1) participates in DNA precursor synthesis and acts as a biomarker for malignant cancer including prostate and breast cancer (*Jagarlamudi, Hansson & Eriksson, 2015*). Meanwhile, research into serological TK1's use in predicting precancer in a study involving 56,178 people showed that serological TK1 protein is a potential proliferative biomarker for early discovery of persons at risk for the development of, or who already have, malignancies or diseases associated with the development of malignancies (*Wang et al., 2018*). Moreover, TK1 is upregulated in the S phase of the cell cycle and its presence in cells is an indicator of active cell proliferation (*Jagarlamudi & Shaw, 2018*). The marker of proliferation, Ki-67 (MKI67), functions to mark tumor cell proliferation, including in the prostate, and has a close relation to the epithelial-mesenchymal transition (EMT) (*Lindsay et al., 2016*). Ki67 may improve the prediction of prostate cancer outcomes based on pathological standard parameters, improving prognosis as well as the monitoring of prostate cancer patients (*Pascale et al., 2016*). TOP2A is considered to be a biomarker for early identification of patients who have increased metastatic potential (*Labbé et al., 2017*). TOP2A encodes for topoisomerase IIα which controls the topology structure of DNA as well as cell cycle progression. This enzyme is a cell proliferation biomarker of cancer and normal tissue that is valuable for prostate cancer treatment (*De Resende et al., 2013; Li et al., 2014*). Cyclin B1 binds to CDC2 to ensure the transition toward mitosis by acting in the cell cycle from the G2 to M phase. High cyclin B1 levels, meanwhile, contribute to the development of polyploidy. Recent research has shown that Cyclin B1 is involved in breast, prostate cancer (*Niranjan et al., 2016; Roh et al., 2005*). A study also showed that

**Wang et al. (2020), *PeerJ*, DOI 10.7717/peerj.8786**

11/21

**Table 4 Functional roles of 20 core genes with degree ≥20.** These biological functions are closely related to the prostate cancer. A total of 20 core genes were screened with degree ≥20. The full name and the function of the 20 core genes were listed.

| No. | Gene symbol | Full name | Function |
|---|---|---|---|
| 1 | CKS2 | Cyclin kinase subunit 2 | Overexpressed CKS2 promotes cell growth and protects cell from apoptosis |
| 2 | TK1 | Thymidine kinase 1 | TK1 is related to DNA precursor synthesis, acts as a proliferating biomarker of prostate and breast cancer |
| 3 | MKI67 | Marker of proliferation Ki-67 | KI-67 is a marker of tumor cell proliferation and the epithelial-mesenchymal transition |
| 4 | TOP2A | Topoisomerase IIα | TOP2A controls the topology structure of DNA and cell cycle progression |
| 5 | CCNB1 | Cyclin B1 | A high cyclin B1 level contributes to the development of polyploidy and was a prognosis biomarker of prostate cancer for chemotherapy |
| 6 | RRM2 | Ribonucleotide-reductase regulatory subunit M2 | RRM2 limits the rate of DNA synthesis and repair. It was considered to be a biomarker to predict recurrence in PCa patients with low-risk |
| 7 | EZH2 | Enhancer of Zeste homolog 2 | Drug targeting therapy EZH2 may be a new therapeutic strategy for advanced PCa and docetaxel-resistant PCa patients |
| 8 | AURKA | Aurora kinase A | AURKA was considered to be a potential prognostic biomarker for the progression of high-risk small-cell prostate cancer |
| 9 | BIRC5 | Baculoviral IAP Repeat containing 5 | BIRC5 may prevent apoptotic cell death, by survivin, a protein that inhibits apoptosis |
| 10 | CCNB2 | Cyclin B2 | CDC2 specifically binds to Cyclin B2 to increase cell migration which is related to the development of CRPC |
| 11 | CDCA8 | Cell division cycle associated 8 | Overexpressed CDCA8 is related to mitosis and cell growth and acts as a prognosis biomarker of breast cancer |
| 12 | TPX2 | Targeting protein for Xenopus kinesin-like protein 2 | TPX2 is a microtubule-associated protein linked to mitosis and spindle assembly and targeting TPX2 is a strategy of PCa |
| 13 | CENPF | Centromere protein F | CENPF acts in the centromere-kinetochore complex and chromosomal segregation, it is related to aggressive prostate cancer |
| 14 | NVAPG | Non-SMC condensin I complex subunit G | NCAPG acts as a target of miR-99a-3p in PCa cells, overexpression of NCAPG is related to CRPC |
| 15 | NUSAP1 | Nucleolar and spindle-associated protein 1 | NUSAP1 is a prognosis biomarker in the early stage of PCa patients |
| 16 | ANLN | Anillin actin binding protein | ANLN was found to be related to cell cycle and growth of PCa cells |
| 17 | CDCA5 | Cell division cycle associated 5 | Knockdown CDCA5 may cause cell cycle arrest in the G2/M phase |
| 18 | ECT2 | Epithelial cell transforming sequence 2 | ECT2 is a guanine nucleotide exchange factor that is related to the progression of tumors |
| 19 | HMMR | Hyaluronan-mediated motility receptor | Hyaluronan-mediated motility receptor (HMMR) binds native and fragmented HA, promotes HA uptake |
| 20 | CENPU | Centromere protein U | CENPU upregulation can increase the invasiveness of prostate cancer |

elevated Cyclin B1 levels served as a biomarker for the prognosis of prostate cancer patients who were being treated with chemotherapy (*Gomez et al., 2007*). Ribonucleotide reductase regulatory subunit M2 (RRM2) is an enzyme that limits the rate of DNA synthesis and repair (*Li et al., 2018*; *Wu et al., 2018*). Mostly, RRM2 was overexpressed in PCa patients with a high Gleason score and an advanced T stage. RRM2 was considered to be a biomarker for prediction of recurrence in low risk PCa patients (*Huang et al., 2014*). In this study, these genes were confirmed to be the core genes in the same protein–protein network. The expression level performed consistently in prostate cancer. By the analysis of GSEA we found that these genes induced prostate cancer consistently by the cell cycle pathway and P53 signaling pathway and GNRH signaling pathway.

Besides the six core genes we mentioned, we found EZH2, AURKA, BIRC5, CCNB2, TPX2, CENPF, NCAPG, NUSAP1 involved in the 20 candidate genes were demonstrated to be risk factors of prostate cancer. Zeste homolog 2 (EZH2) acts as the methyltransferase component of PRC2. Its enhancer regulation disorder is widely found in many aggressive, advanced cancers (*Murashima et al., 2019*). PRC2 inhibits stem cell self-renewal, cell cycle, cell differentiation and cell transformation through EZH2 (H3K27me3) modification (*Alzrigat, Jernberg-Wiklund & Licht, 2018*). Drug therapies targeting EZH2 may be a new strategy for advanced PCa and docetaxel-resistant PCa patients (*Liu et al., 2019*; *Qiu et al., 2019*). Aurora kinase A is encoded by the AURKA gene and has a vital function in the development of the cell cycle. It both controls and promotes entry into mitosis (*López-Cortés et al., 2018*). Studies have shown that AURKA is linked to pathological stage and distant metastasis in HCC (*Chen et al., 2017*). The AURKA gene has been proven to amplify in 67% of PCa patients with highly aggressive hormone-naive castration resistant cancer. AURKA, therefore, was considered to be a potential prognostic biomarker for the progression of high-risk small-cell prostate cancer's resistance to castration (*Park et al., 2014*). One aspect of cancer is that apoptosis is uninhibited. Survivin, an inhibitor-of-apoptosis protein, is encoded by BIRC5, a gene that is linked to the regulation of apoptosis and cell division (*Moore et al., 2014*; *Hmeljak et al., 2011*). Studies have shown that BIRC5 is a biomarker for Oral Squamous Cell Carcinoma as well as breast, liver and prostate cancer. In PTEN deletion mouse model, there was a positive correlation between the survivin level and tumor growth. Researchers found that survivin plays an important role in the conversion process of prostatic intraepithelial neoplasia to adenocarcinoma (*Adisetiyo et al., 2013*) CCNB2 is linked to the process of transition from the G2 to the M phase. It has acted as a prognosis biomarker of non-small-cell lung cancer (*Qian et al., 2015*). Recently, studies have shown that CDC2 specifically binds to Cyclin B2 to increase cell migration, relating to development in CRPC (*Huang et al., 2017*; *Manes et al., 2003*). The targeting protein for Xenopus kinesin-like protein 2 (TPX2) is a microtubule associated protein that targets TPX2 repressed breast cancer by inhibiting the PI3k/AKT/P21 signaling pathway and activating the p53 pathway (*Chen et al., 2018*). High levels of chromosome missegregation is related to cell death and tumorigenesis. Recently, studies have shown that targeting TPX2 in breast and prostate cancer lowered the rate of chromosome missegregation, and have therefore regarded TPX2 as a candidate biomarker for treatment (*Pan et al., 2017*). Thus, TPX2 is considered to be a candidate target for PCa patients. Centromere protein F participates in cancer metabolism by regulating pyruvate kinase M2 phosphorylation signaling (*Shahid et al., 2018*). Recent work suggests that the upregulation of CENPF is linked to aggressive prostate cancer (*Göbel et al., 2018*). NCAPG has been demonstrated to be related to the overexpression of CCNB1. It is suggested to be a candidate target for HCC treatment (*Zhang et al., 2018*). The overexpression of Non-SMC condensing I complex subunit G (NCAPG) is involved in CRPC, and thus it may be a biomarker for PCa (*Arai et al., 2018*). Nucleolar and spindle-associated protein 1 (NUSAP1) is a prognosis biomarker in the in the earliest stage of PCa (*Gordon et al., 2017*). The overexpression of

NUSAP1 may be related to the increased invasion and proliferation of PCa cells through the loss of RB1 (*Gordon, Gulzar & Brooks, 2015*).

Afterwards, we found the researches of CDCA8, ANLN, CDCA5, ECT2, HMMR, CENPU were poor, which the prognosis value and mechanisms needed further investigation. Cell division cycle associated 8 (CDCA8) over-expression is related to mitosis and tumor growth and may act as a prognosis biomarker in bladder cancer, cutaneous melanoma, breast cancer and osteosarcoma (*Bi et al., 2018*; *Dai et al., 2015*). The regulation mechanism of CDCA8 in PCa patients, however, is unclear and needed further study. ANLN was found to relate to the cell cycle and growth in PCa, but the regulation mechanisms are currently unknown (*Takayama et al., 2019*). CDCA5 knockdown led to cell cycle arrest in the G2/M phase (*Tian et al., 2018*). Epithelial cell transforming sequence 2 (ECT2) acts as an exchange factor of the guanine nucleotide, which is related to the progression of cell division regulation and the cell cycle (*Bai et al., 2018*). Hyaluronan-mediated motility receptor (HMMR) promotes HA uptake, and related time to biochemical failure in Gleason score 7 tumor (*Rizzardi et al., 2014*). Overexpression of CENPU is related to breast cancer, lung cancer, ovarian cancer and prostate cancer. Upregulated CENPU can increase the invasiveness of PCa cells (*Winter et al., 2017*).

In this study, we identified CKS2, TK1, MKI67, TOP2A, CCNB1 and RRM2 as crucial components in the diagnosis and treatment of PCa. They acted together in the same pathway to induce prostate cancer. The research of the correlation between molecules in cell cycle, DNA replication, GNRH signaling pathway, P53_signaling pathway could reveal the underlying causes of cancer and provide novel ideas for research into target drugs. Meanwhile, the expression difference of CDCA8, ANLN, CDCA5, ECT2, HMMR and CENPU were significantly which needed in-deep study.

## CONCLUSION

This study excavated the core genes of prostate cancer, analyzed their functions, pathways, and their phenotype by means of reliable bioinformatics analysis of multiple datasets. The core genes in this study were considered to be potential targets and biomarkers, providing new ideas for the diagnosis and treatment of prostate cancer. More experimental studies are needed, however, to verify the mechanisms of these genes in prostate cancer.

## ADDITIONAL INFORMATION AND DECLARATIONS

## Competing Interests

The authors declare that they have no competing interests.

## Author Contributions

- Yutao Wang performed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Jianfeng Wang conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Kexin Yan analyzed the data, prepared figures and/or tables, and approved the final draft.
- Jiaxing Lin performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Zhenhua Zheng analyzed the data, prepared figures and/or tables, and approved the final draft.
- Jianbin Bi conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

## Data Availability

The following information was supplied regarding data availability:

Data is available at NCBI GEO: GSE38241, GSE69223, GSE46602, GSE104749.

## REFERENCES

**Adisetiyo H, Liang M, Liao CP, Aycock-Williams A, Cohen MB, Xu S, Neamati N, Conway EM, Cheng CY, Nikitin AY, Roy-Burman P. 2013.** Loss of survivin in the prostate epithelium impedes carcinogenesis in a mouse model of prostate adenocarcinoma. *PLOS ONE* **8**:e69484.

**Alzrigat M, Jernberg-Wiklund H, Licht JD. 2018.** Targeting EZH2 in multiple myeloma-multifaceted anti-tumor activity. *Epigenomes* **2(3)**:16 DOI 10.3390/epigenomes2030016.

**Arai T, Okato A, Yamada Y, Sugawara S, Kurozumi A, Kojima S, Yamazaki K, Naya Y, Ichikawa T, Seki N. 2018.** Regulation of NCAPG by miR-99a-3p (passenger strand) inhibits cancer cell aggressiveness and is involved in CRPC. *Cancer Medicine* **7(5)**:1988–2002 DOI 10.1002/cam4.1455.

**Aryee MJ, Liu W, Engelmann JC, Nuhn P, Gurel M, Haffner MC, Esopi D, Irizarry RA, Getzenberg RH, Nelson WG, Luo J, Xu J, Isaacs WB, Bova GS, Yegnasubramanian S. 2013.** DNA methylation alterations exhibit intraindividual stability and interindividual heterogeneity in prostate cancer metastases. *Science Translational Medicine* **5(169)**:169ra10 DOI 10.1126/scitranslmed.3005211.

**Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000.** Gene ontology: tool for the unification of biology the gene ontology consortium. *Nature Genetics* **25**:25–29.

**Bai X, Yi M, Xia X, Yu S, Zheng X, Wu K. 2018.** Progression and prognostic value of ECT2 in non-small-cell lung cancer and its correlation with PCNA. *Cancer Management and Research* **10**:4039–4050 DOI 10.2147/CMAR.S170033.

**Bandettini WP, Kellman P, Mancini C, Booker OJ, Vasu S, Leung SW, Wilson JR, Shanbhag SM, Chen MY, Arai AE. 2012.** MultiContrast delayed enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. *Journal of Cardiovascular Magnetic Resonance* **14(1)**:83 DOI 10.1186/1532-429X-14-83.

**Bi Y, Chen S, Jiang J, Yao J, Wang G, Zhou Q, Li S. 2018.** CDCA8 expression and its clinical relevance in patients with bladder cancer. *Medicine* **97(34)**:e11899 DOI 10.1097/MD.0000000000011899.

**Cancer Genome Atlas Research Network TCGA. 2015.** The molecular taxonomy of primary prostate cancer. *Cell* **163(4)**:1011–1025.

**Chen C, Song G, Xiang J, Zhang H, Zhao S, Zhan Y. 2017.** AURKA promotes cancer metastasis by regulating epithelial-mesenchymal transition and cancer stem cell properties in hepatocellular carcinoma. *Biochemical and Biophysical Research Communications* **486(2)**:514–520 DOI 10.1016/j.bbrc.2017.03.075.

**Chen M, Zhang H, Zhang G, Zhong A, Ma Q, Kai J, Tong Y, Xie S, Wang Y, Zheng H, Guo L, Lu R. 2018.** Targeting TPX2 suppresses proliferation and promotes apoptosis via repression of the PI3k/AKT/P21 signaling pathway and activation of p53 pathway in breast cancer. *Biochemical and Biophysical Research Communications* **507(1–4)**:74–82 DOI 10.1016/j.bbrc.2018.10.164.

**Dai C, Miao C-X, Xu X-M, Liu L-J, Gu Y-F, Zhou D, Chen L-S, Lin G, Lu G-X. 2015.** Transcriptional activation of human CDCA8 gene regulated by transcription factor NF-Y in embryonic stem cells and cancer cells. *Journal of Biological Chemistry* **290(37)**:22423–22434 DOI 10.1074/jbc.M115.642710.

**De Resende MF, Vieira S, Chinen LT, Chiappelli F, Da Fonseca FP, Guimarães GC, Soares FA, Neves I, Pagotty S, Pellionisz PA, Barkhordarian A, Brant X, Rocha RM. 2013.** Prognostication of prostate cancer based on TOP2A protein and gene assessment: TOP2A in prostate cancer. *Journal of Translational Medicine* **11(1)**:36 DOI 10.1186/1479-5876-11-36.

**Edgar R, Domrachev M, Lash AE. 2002.** Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research* **30(1)**:207–210 DOI 10.1093/nar/30.1.207.

**Farhood B, Mortezaee K, Haghi-Aminjan H, Khanlarkhani N, Salehi E, Nashtaei MS, Najafi M, Sahebkar A. 2019.** A systematic review of radiation-induced testicular toxicities following radiotherapy for prostate cancer. *Journal of Cellular Physiology* **234(9)**:14828–14837 DOI 10.1002/jcp.28283.

**Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, Von Mering C, Jensen LJ. 2013.** STRING v9.1: protein–protein interaction networks, with increased coverage and integration. *Nucleic Acids Research* **41(D1)**:D808–D815 DOI 10.1093/nar/gks1094.

**Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N. 2013.** Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science Signaling* **6(269)**:pl1 DOI 10.1126/scisignal.2004088.

**Gomez LA, De Las Pozas A, Reiner T, Burnstein K, Perez-Stable C. 2007.** Increased expression of cyclin B1 sensitizes prostate cancer cells to apoptosis induced by chemotherapy. *Molecular Cancer Therapeutics* **6(5)**:1534–1543 DOI 10.1158/1535-7163.MCT-06-0727.

**Gordon CA, Gulzar ZG, Brooks JD. 2015.** *NUSAP1* expression is upregulated by loss of RB1 in prostate cancer cells. *Prostate* **75(5)**:517–526 DOI 10.1002/pros.22938.

**Gordon CA, Gong X, Ganesh D, Brooks JD. 2017.** NUSAP1 promotes invasion and metastasis of prostate cancer. *Oncotarget* **8**:29935–29950.

**Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, Quist MJ, Jing X, Lonigro RJ, Brenner JC, Asangani IA, Ateeq B, Chun SY, Siddiqui J, Sam L, Anstett M, Mehra R, Prensner JR, Palanisamy N, Ryslik GA, Vandin F, Raphael BJ, Kunju LP, Rhodes DR, Pienta KJ, Chinnaiyan AM, Tomlins SA. 2012.** The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487(7406)**:239–243 DOI 10.1038/nature11125.

**Göbel C, Özden C, Schroeder C, Hube-Magg C, Kluth M, Möller-Koop C, Neubauer E, Hinsch A, Jacobsen F, Simon R, Sauter G, Michl U, Pehrke D, Huland H, Graefen M, Schlomm T, Luebke AM. 2018.** Upregulation of centromere protein F is linked to aggressive prostate cancers. *Cancer Management and Research* **10**:5491–5504 DOI 10.2147/CMAR.S165630.

**Haeussler M, Zweig AS, Tyner C, Speir ML, Rosenbloom KR, Raney BJ, Lee CM, Lee BT, Hinrichs AS, Gonzalez JN, Gibson D, Diekhans M, Clawson H, Casper J, Barber GP, Haussler D, Kuhn RM, Kent WJ. 2019.** The UCSC genome browser database: 2019 update. *Nucleic Acids Research* **47(D1)**:D853–D858 DOI 10.1093/nar/gky1095.

**Hemminki K. 2012.** Familial risk and familial survival in prostate cancer. *World Journal of Urology* **30(2)**:143–148 DOI 10.1007/s00345-011-0801-1.

**Hernández G, Ramírez JL, Pedroza-Torres A, Herrera LA, Jiménez-Ríos MA. 2019.** The secret life of translation initiation in prostate cancer. *Frontiers in Genetics* **10**:1415 DOI 10.3389/fgene.2019.00014.

**Hmeljak J, Erčulj N, Dolžan V, Kern I, Cör A. 2011.** BIRC5 promoter SNPs do not affect nuclear survivin expression and survival of malignant pleural mesothelioma patients. *Journal of Cancer Research and Clinical Oncology* **137(11)**:1641–1651 DOI 10.1007/s00432-011-1030-0.

**Huang CG, Li FX, Pan S, Xu CB, Dai JQ, Zhao XH. 2017.** Identification of genes associated with castration-resistant prostate cancer by gene expression profile analysis. *Molecular Medicine Reports* **16(5)**:6803–6813 DOI 10.3892/mmr.2017.7488.

**Huang Y, Liu X, Wang Y-H, Yeh S-D, Chen C-L, Nelson RA, Chu P, Wilson T, Yen Y. 2014.** The prognostic value of ribonucleotide reductase small subunit M2 in predicting recurrence for prostate cancers. *Urologic Oncology: Seminars and Original Investigations* **32(1)**:51.e9–51.e19 DOI 10.1016/j.urolonc.2013.08.002.

**Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, Stephens R, Baseler MW, Lane HC, Lempicki RA. 2007.** The DAVID gene functional classification tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biology* **8(9)**:R183 DOI 10.1186/gb-2007-8-9-r183.

**Islami F, Moreira DM, Boffetta P, Freedland SJ. 2014.** A systematic review and meta-analysis of tobacco use and prostate cancer mortality and incidence in prospective cohort studies. *European Urology* **66(6)**:1054–1064 DOI 10.1016/j.eururo.2014.08.059.

**Ito K, Murphy D. 2013.** Application of ggplot2 to pharmacometric graphics. *CPT Pharmacometrics and Systems Pharmacology* **2**:e79 DOI 10.1038/psp.2013.56.

**Jagarlamudi KK, Hansson LO, Eriksson S. 2015.** Breast and prostate cancer patients differ significantly in their serum Thymidine kinase 1 (TK1) specific activities compared with those hematological malignancies and blood donors: implications of using serum TK1 as a biomarker. *BMC Cancer* **15(1)**:66 DOI 10.1186/s12885-015-1073-8.

**Jagarlamudi KK, Shaw M. 2018.** Thymidine kinase 1 as a tumor biomarker: technical advances offer new potential to an old biomarker. *Biomarkers in Medicine* **12(9)**:1035–1048 DOI 10.2217/bmm-2018-0157.

**Jansson KF, Akre O, Garmo H, Bill-Axelson A, Adolfsson J, Stattin P, Bratt O. 2012.**
Concordance of tumor differentiation among brothers with prostate cancer. *European Urology*
**62(4)**:656–661 DOI 10.1016/j.eururo.2012.02.032.

**Ju-Kun S, Yuan D-B, Rao H-F, Chen T-F, Luan B-S, Xu X-M, Jiang F-N, Zhong W-D, Zhu J-G.
2016.** Association between Cd exposure and risk of prostate cancer: a PRISMA-compliant
systematic review and meta-analysis. *Medicine* **95(6)**:e2708
DOI 10.1097/MD.0000000000002708.

**Kanehisa M. 2002.** The KEGG database. *Novartis Foundation Symposium* **247**:91–252.

**Katarzyna T, Patrycja C, Maciej W. 2015.** The cancer genome atlas (TCGA): an immeasurable
source of knowledge. *Contemporary Oncology* **19(1A)**:68–77 DOI 10.5114/wo.2014.47136.

**Labbé DP, Sweeney CJ, Brown M, Galbo P, Rosario S, Wadosky KM, Ku SY, Sjöström M,
Alshalalfa M, Erho N, Davicioni E, Karnes RJ, Schaeffer EM, Jenkins RB, Den RB, Ross AE,
Bowden M, Huang Y, Gray KP, Feng FY, Spratt DE, Goodrich DW, Eng KH, Ellis L. 2017.**
TOP2A and EZH2 provide early detection of an aggressive prostate cancer subgroup.
*Clinical Cancer Research* **23(22)**:7072–7083 DOI 10.1158/1078-0432.CCR-17-0413.

**Lacoste J, Aprikian AG, Chevalier S. 2005.** Focal adhesion kinase is required for
bombesin-induced prostate cancer cell motility. *Molecular and Cellular Endocrinology*
**235(1–2)**:51–61 DOI 10.1016/j.mce.2004.06.014.

**Lan Y, Zhang Y, Wang J, Lin C, Ittmann MM, Wang F. 2008.** Aberrant expression of Cks1
and Cks2 contributes to prostate tumorigenesis by promoting proliferation and inhibiting
programmed cell death. *International Journal of Cancer* **123(3)**:543–551 DOI 10.1002/ijc.23548.

**Li L, Lei Q, Zhang S, Kong L, Qin B. 2017.** Screening and identification of key biomarkers in
hepatocellular carcinoma: evidence from bioinformatic analysis. *Oncology Reports*
**38(5)**:2607–2618 DOI 10.3892/or.2017.5946.

**Li X, Liu Y, Chen W, Fang Y, Xu H, Zhu HH, Chu M, Li W, Zhuang G, Gao WQ. 2014.** TOP2A
high is the phenotype of recurrence and metastasis whereas TOP2A neg cells represent cancer
stem cells in prostate cancer. *Oncotarget* **5**:9498–9513.

**Li J, Pang J, Liu Y, Zhang J, Zhang C, Shen G, Song L. 2018.** Suppression of RRM2 inhibits cell
proliferation, causes cell cycle arrest and promotes the apoptosis of human neuroblastoma cells
and in human neuroblastoma RRM2 is suppressed following chemotherapy. *Oncology Reports*
**40**:355–360 DOI 10.3892/or.2018.6420.

**Lian W-Q, Luo F, Song X-L, Lu Y-J, Zhao S-C. 2015.** Gonorrhea and prostate cancer incidence: an
updated meta-analysis of 21 epidemiologic studies. *Medical Science Monitor* **21**:1895–1903
DOI 10.12659/MSM.893579.

**Lindsay CR, Le Moulec S, Billiot F, Loriot Y, Ngo-Camus M, Vielh P, Fizazi K, Massard C,
Farace F. 2016.** Vimentin and Ki67 expression in circulating tumor cells derived from
castrate-resistant prostate cancer. *BMC Cancer* **16(1)**:168 DOI 10.1186/s12885-016-2192-6.

**Liu Q, Wang G, Li Q, Jiang W, Kim J-S, Wang R, Zhu S, Wang X, Yan L, Yi Y, Zhang L,
Meng Q, Li C, Zhao D, Qiao Y, Li Y, Gursel D-B, Chinnaiyan A-M, Chen K, Cao Q. 2019.**
Polycomb group proteins EZH2 and EED directly regulate androgen receptor in
advanced prostate cancer. *International Journal of Cancer* **145(2)**:415–426
DOI 10.1002/ijc.32118.

**López-Cortés A, Cabrera-Andrade A, Oña-Cisneros F, Echeverría C, Rosales F, Ortiz M,
Tejera E, Paz-Y-Miño C. 2018.** Breast cancer risk associated with genotype polymorphisms
of the aurora kinase a gene (AURKA): a case-control study in a high altitude ecuadorian
mestizo population. *Pathology & Oncology Research* **24(3)**:457–465
DOI 10.1007/s12253-017-0267-6.

Manes T, Zheng D-Q, Tognin S, Woodard AS, Marchisio PC, Languino LR. 2003. $\alpha_v\beta_3$ integrin expression up-regulates cdc2, which modulates cell migration. *Journal of Cell Biology* **161(4)**:817–826 DOI 10.1083/jcb.200212172.

Mayer R, Klemen H, Quehenberger F, Sankin O, Mayer E, Hackl A, Smolle-Juettner F-M. 2001. Hyperbaric oxygen—an effective tool to treat radiation morbidity in prostate cancer. *Radiotherapy and Oncology* **61(2)**:151–156 DOI 10.1016/S0167-8140(01)00430-3.

Meller S, Meyer H-A, Bethan B, Dietrich D, Maldonado SG, Lein M, Montani M, Reszka R, Schatz P, Peter E, Stephan C, Jung K, Kamlage B, Kristiansen G. 2016. Integration of tissue metabolomics, transcriptomics and immunohistochemistry reveals ERG- and gleason score-specific metabolomic alterations in prostate cancer. *Oncotarget* **7(2)**:1421–1438 DOI 10.18632/oncotarget.6370.

Moore AS, Alonzo TA, Gerbing RB, Lange BJ, Heerema NA, Franklin J, Raimondi SC, Hirsch BA, Gamis AS, Meshinchi S. 2014. *BIRC5* (survivin) splice variant expression correlates with refractory disease and poor outcome in pediatric acute myeloid leukemia: a report from the children's oncology group. *Pediatric Blood & Cancer* **61(4)**:647–652 DOI 10.1002/pbc.24822.

Mortensen MM, Høyer S, Lynnerup A-S, Ørntoft TF, Sørensen KD, Borre M, Dyrskjøt L. 2015. Expression profiling of prostate cancer tissue delineates genes associated with recurrence after prostatectomy. *Scientific Reports* **5**:16018 DOI 10.1038/srep16018.

Murashima A, Shinjo K, Katsushima K, Onuki T, Kondoh Y, Osada H, Kagaya N, Shin-Ya K, Kimura H, Yoshida M, Murakami S, Kondo Y. 2019. Identification of a chemical modulator of EZH2-mediated silencing by cell-based high-throughput screening assay. *Journal of Biochemistry* **166(1)**:41–50 DOI 10.1093/jb/mvz007.

Niranjan KC, Tayaar A, Kumar GS, Krishnapillai R, Hallikeri K, Hunasgi S. 2016. Immunohistochemical expression of cyclin B1 in epithelial hyperplasia, dysplasia and oral squamous cell carcinomas: a comparative study. *Journal of Clinical and Diagnostic Research* **10**:ZC85–ZC90 DOI 10.7860/JCDR/2016/19820.8563.

Pan H-W, Su H-H, Hsu C-W, Huang G-J, Wu TT-L. 2017. Targeted TPX2 increases chromosome missegregation and suppresses tumor cell growth in human prostate cancer. *OncoTargets and Therapy* **10**:3531–3543 DOI 10.2147/OTT.S136491.

Park K, Chen Z, MacDonald TY, Siddiqui J, Ye H, Erbersdobler A, Shevchuk MM, Robinson BD, Sanda MG, Chinnaiyan AM, Beltran H, Rubin MA, Mosquera JM. 2014. Prostate cancer with paneth cell-like neuroendocrine differentiation has recognizable histomorphology and harbors AURKA gene amplification. *Human Pathology* **45(10)**:2136–2143 DOI 10.1016/j.humpath.2014.06.008.

Pascale M, Aversa C, Barbazza R, Marongiu B, Siracusano S, Stoffel F, Sulfaro S, Roggero E, Bonin S, Stanta G. 2016. The proliferation marker Ki67, but not neuroendocrine expression, is an independent factor in the prediction of prognosis of primary prostate cancer patients. *Radiology and Oncology* **50(3)**:313–320 DOI 10.1515/raon-2016-0033.

Qian X, Song X, He Y, Yang Z, Sun T, Wang J, Zhu G, Xing W, You C. 2015. CCNB2 overexpression is a poor prognostic biomarker in Chinese NSCLC patients. *Biomedicine & Pharmacotherapy* **74**:222–227 DOI 10.1016/j.biopha.2015.08.004.

Qiu X, Wang W, Li B, Cheng B, Lin K, Bai J, Li H, Yang G. 2019. Targeting Ezh2 could overcome docetaxel resistance in prostate cancer cells. *BMC Cancer* **19**:27 DOI 10.1186/s12885-018-5228-2.

**Rao D, Yu H, Bai Y, Zheng X, Xie L. 2015.** Does night-shift work increase the risk of prostate cancer? A systematic review and meta-analysis. *Onco Targets and Therapy* **8**:2817–2826 DOI 10.2147/OTT.S89769.

**Rizzardi AE, Vogel RI, Koopmeiners JS, Forster CL, Marston LO, Rosener NK, Akentieva N, Price MA, Metzger GJ, Warlick CA, Henriksen JC, Turley EA, McCarthy JB, Schmechel SC. 2014.** Elevated hyaluronan and hyaluronan-mediated motility receptor are associated with biochemical failure in patients with intermediate-grade prostate tumors. *Cancer* **120(12)**:1800–1809 DOI 10.1002/cncr.28646.

**Roh M, Song C, Kim J, Abdulkadir SA. 2005.** Chromosomal instability induced by Pim-1 is passage-dependent and associated with dysregulation of cyclin B1. *Journal of Biological Chemistry* **280(49)**:40568–40577 DOI 10.1074/jbc.M509369200.

**Shahid M, Lee MY, Piplani H, Andres AM, Zhou B, Yeon A, Kim M, Kim HL, Kim J. 2018.** Centromere protein F (CENPF), a microtubule binding protein, modulates cancer metabolism by regulating pyruvate kinase M2 phosphorylation signaling. *Cell Cycle* **17(24)**:2802–2818 DOI 10.1080/15384101.2018.1557496.

**Shan M, Xia Q, Yan D, Zhu Y, Zhang X, Zhang G, Guo J, Hou J, Chen W, Zhu T, Zhang X, Xu J, Wang J, Ding T, Zheng J. 2017.** Molecular analyses of prostate tumors for diagnosis of malignancy on fine-needle aspiration biopsies. *Oncotarget* **8**:104761–104771.

**Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. 2011.** New features for data integration and network visualization. *Bioinformatics* **27(3)**:431–432 DOI 10.1093/bioinformatics/btq675.

**Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005.** Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102(43)**:15545–15550 DOI 10.1073/pnas.0506580102.

**Tai K-Y, Shiah S-G, Shieh Y-S, Kao Y-R, Chi C-Y, Huang E, Lee H-S, Chang L-C, Yang P-C, Wu C-W. 2007.** DNA methylation and histone modification regulate silencing of epithelial cell adhesion molecule for tumor invasion and progression. *Oncogene* **26(27)**:3989–3997 DOI 10.1038/sj.onc.1210176.

**Takayama K-I, Suzuki Y, Yamamoto S, Obinata D, Takahashi S, Inoue S. 2019.** Integrative genomic analysis of OCT1 reveals coordinated regulation of androgen receptor in advanced prostate cancer. *Endocrinology* **160(2)**:463–472 DOI 10.1210/en.2018-00923.

**Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, Arora VK, Kaushik P, Cerami E, Reva B, Antipin Y, Mitsiades N, Landers T, Dolgalev I, Major JE, Wilson M, Socci ND, Lash AE, Heguy A, Eastham JA, Scher HI, Reuter VE, Scardino PT, Sander C, Sawyers CL, Gerald WL. 2010.** Integrative genomic profiling of human prostate cancer. *Cancer Cell* **18(1)**:11–22 DOI 10.1016/j.ccr.2010.05.026.

**Tian Y, Wu J, Chagas C, Du Y, Lyu H, He Y, Qi S, Peng Y, Hu J. 2018.** CDCA5 overexpression is an indicator of poor prognosis in patients with hepatocellular carcinoma (HCC). *BMC Cancer* **18(1)**:1187 DOI 10.1186/s12885-018-5072-4.

**Vanaja DK, Cheville JC, Iturria SJ, Young CY. 2003.** Transcriptional silencing of zinc finger protein 185 identified by expression profiling is associated with prostate cancer progression. *Cancer Research* **63**:3877–3882.

**Varambally S, Yu J, Laxman B, Rhodes DR, Mehra R, Tomlins SA, Shah RB, Chandran U, Monzon FA, Becich MJ, Wei JT, Pienta KJ, Ghosh D, Rubin MA, Chinnaiyan AM. 2005.** Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. *Cancer Cell* **8(5)**:393–406 DOI 10.1016/j.ccr.2005.10.001.

**Wang Y, Jiang X, Wang S, Yu H, Zhang T, Xu S, Li W, He E, Skog S. 2018.** Serological TK1 predict pre-cancer in routine health screenings of 56,178 people. *Cancer Biomarkers* **22(2)**:237–247 DOI 10.3233/CBM-170846.

**Winter JM, Gildea DE, Andreas JP, Gatti DM, Williams KA, Lee M, Hu Y, Zhang S, NISC Comparative Sequencing Program, Mullikin JC, Wolfsberg TG, McDonnell SK, Fogarty ZC, Larson MC, French AJ, Schaid DJ, Thibodeau SN, Churchill GA, Crawford NP. 2017.** Mapping complex traits in a diversity outbred F1 mouse population identifies germline modifiers of metastasis in human prostate cancer. *Cell Systems* **4**:31–45.e6.

**Wu Y, Sun J, Li A, Chen D. 2018.** The promoted delivery of RRM2 siRNA to vascular smooth muscle cells through liposome-polycation-DNA complex conjugated with cell penetrating peptides. *Biomedicine & Pharmacotherapy* **103**:982–988 DOI 10.1016/j.biopha.2018.03.068.

**Yang Y, Jia B, Zhao X, Wang Y, Ye W. 2018.** miR-93-5p may be an important oncogene in prostate cancer by bioinformatics analysis. *Journal of Cellular Biochemistry* **120(6)**:10463–10483 DOI 10.1002/jcb.28332.

**Yu M, Zhong M, Qiao Z. 2013.** Expression and clinical significance of cyclin kinase subunit 2 in colorectal cancer. *Oncology Letters* **6(3)**:777–780 DOI 10.3892/ol.2013.1456.

**Zhang Q, Su R, Shan C, Gao C, Wu P. 2018.** Non-SMC condensin I complex, subunit G (NCAPG) is a novel mitotic gene required for hepatocellular cancer cell proliferation and migration. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics* **26(2)**:269–276 DOI 10.3727/096504017X15075967560980.

**Wang et al. (2020), *PeerJ*, DOI 10.7717/peerj.8786**

21/21