

RESEARCH ARTICLE

Identification of UBE2C as hub gene in driving prostate cancer by integrated bioinformatics analysis

Yan Wang^{1,2}, Jili Wang^{1,2}, Qiusu Tang¹, Guoping Ren^{1,2*}

1 Department of Pathology, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, **2** Department of Pathology and Pathophysiology, Zhejiang University School of Medicine, Hangzhou, China

* gpren2002@163.com

Abstract

Background

The aim of this study was to identify novel genes in promoting primary prostate cancer (PCa) progression and to explore its role in the prognosis of prostate cancer.

Methods

Four microarray datasets containing primary prostate cancer samples and benign prostate samples were downloaded from Gene Expression Omnibus (GEO), then differentially expressed genes (DEGs) were identified by R software (version 3.6.2). Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were performed to identify the function of DEGs. Using STRING and Cytoscape (version 3.7.1), we constructed a protein-protein interaction (PPI) network and identified the hub gene of prostate cancer. Clinical data on GSE70770 and TCGA was collected to show the role of hub gene in prostate cancer progression. The correlations between hub gene and clinical parameters were also indicated by cox regression analysis. Gene Set Enrichment Analysis (GSEA) was performed to highlight the function of Ubiquitin-conjugating enzyme complex (UBE2C) in prostate cancer.

Results

243 upregulated genes and 298 downregulated genes that changed in at least two microarrays have been identified. GO and KEGG analysis indicated significant changes in the oxidation-reduction process, angiogenesis, TGF-beta signaling pathway. UBE2C, PDZ-binding kinase (PBK), cyclin B1 (CCNB1), Cyclin-dependent kinase inhibitor 3 (CDKN3), topoisomerase II alpha (TOP2A), Aurora kinase A (AURKA) and MKI67 were identified as the candidate hub genes, which were all correlated with prostate cancer patient' disease-free survival in TCGA. In fact, only UBE2C was highly expressed in prostate cancer when compared with benign prostate tissue in TCGA and the expression of UBE2C was also in parallel with the Gleason score of prostate cancer. Cox regression analysis has indicated UBE2C could function as the independent prognostic factor of prostate cancer. GSEA

OPEN ACCESS

Citation: Wang Y, Wang J, Tang Q, Ren G (2021) Identification of UBE2C as hub gene in driving prostate cancer by integrated bioinformatics analysis. PLoS ONE 16(2): e0247827. <https://doi.org/10.1371/journal.pone.0247827>

Editor: Surinder K. Batra, University of Nebraska Medical Center, UNITED STATES

Received: August 22, 2020

Accepted: February 14, 2021

Published: February 25, 2021

Copyright: © 2021 Wang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are available from the NCBI Gene Expression Omnibus (accession numbers GSE3325, GSE46602, GSE70770, GSE69223, GSE104749, GSE116918).

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

showed UBE2C had played an important role in the pathway of prostate cancer, such as NOTCH signaling pathway, WNT- β -catenin signaling pathway.

Conclusions

UBE2C was pivotal for the progression of prostate cancer and the level of UBE2C was important to predict the prognosis of patients.

Introduction

Prostate cancer has become the most universal diagnosed cancer among American men, which accounted for 20% of newly diagnosed cancers [1,2]. To our surprise, although the survival probability of prostate cancer was the highest, the proleptic mortality of prostate carcinoma accounted for 10% among the whole cancer death cases [1]. Among American men, prostate cancer remained the third-leading reason for cancer death. Based on the cancer statistics in China, we have concluded that prostate cancer has become the most popular cancer of the genitourinary system of men [2]. In detail, from 2000 to 2011, the occurrence of prostate cancer has increased year by year. Prostate cancer has become a threat to long-term health of patients and has worsened the living conditions. Unfortunately, in China, the mortality of prostate cancer has been rising year by year too. It was essential for clinicians to diagnose and to treat diverse kinds of prostate cancer correctly and appropriately. For localized prostate cancer, surgery and radiation have become the first choice, although it could bring some adverse effects to patients, such as frequent urination, hematuria and sexual dysfunction that maybe pull down life quality [3]. Androgen deprivation therapy (ADT) combined with radiotherapy was superior to other therapies for high risk prostate cancer patients [4]. Although ADT was recognized as standard therapy for diverse prostate cancer, chemotherapy has also been recommended for metastatic prostate cancer. In terms of the screening and diagnosis, prostate-specific antigen (PSA) has been widely performed to monitor and diagnose prostate cancer. While due to PSA's poor specificity, early PSA test, could lead to overdiagnosis and overtreatment of prostate cancer [5]. Previous studies have proved high-sequencing could function as a novel way to identify biomarkers in various cancers. Fang et al. have compared stroma surrounding invasive prostate tumors and matched normal stroma, in order to identify a potential target for prostate cancer [6]. Comparing luminal cells, basal cells and epithelial cells from hormone-naive and castration-resistant mice, L6YD was identified as a marker castration-resistant prostate cancer [7]. Thus, this article aimed to identify the hub genes of prostate carcinoma by bioinformatics analysis.

In this study, four databases were downloaded from GEO, which included benign prostate gland and primary prostate cancer tissue. Using limma and affy package of R software, DEGs were identified. Furthermore, GO and KEGG pathway analysis were performed by The Database for Annotation, Visualization, and Integrated Discovery (DAVID) [8] (<http://www.david.niaid.nih.gov>). Then, PPI and modules analysis was performed by STRING [9] (<https://string-db.org/>), metaspape [10] (<http://metaspape.org/gp/index.html>) and Cytoscape [11,12]. Survival analysis was performed to screen and validate the hub gene of prostate cancer. In brief, UBE2C, highly expressed in prostate cancer tissue, was identified as the real hub gene. Through GSEA [13] and univariate and multivariate analysis, the role of UBE2C in prostate cancer was explored completely.

Materials and methods

Microarray data

GSE104749 [14], GSE3325 [15], GSE69223 [16], GSE46602 [17], four prostate cancer gene expression profiles were downloaded from GEO, which were from the same GPL platform (GPL570). The criteria of inclusion GEO gene expression profiles were shown in S1 Fig. GSE3325 contained 7 localized prostate cancer tissues and 6 benign prostate tissues. GSE46602 contained 36 prostate tumor samples and 14 normal prostate gland samples. GSE69223 included 15 cancer tissues and 15 corresponding noncancerous tissue samples. GSE104749 contained 4 prostate cancer samples and 4 benign prostate gland samples (Table 1). The clinicopathological features of 206 primary prostate cancer patients in the GSE70770 were collected for further analysis.

Identification of DEGs

GSE3325, GSE46602, GSE69223 and GSE104749 gene expression profiles have been normalized and transformed into suitable matrix using Robust Multichip Average (RMA) included in affy package of R software (version 3.6.1) consisting of background adjustment, quantile normalization and summarization pre-processing [6,18]. Limma package was used to identify DEGs between primary prostate cancer and noncancerous prostate samples. The criteria of DEGs were adjusted p value < 0.05 and $|\log_2(\text{fold change})| > 1$. Venn diagram was used to find DEGs appearing in at least two databases (Figs 1 and S2).

GO and KEGG pathway analysis of the DEGs

To further explore the function of DEGs, GO [19] and KEGG [20] analysis of DEGs was performed by DAVID. We downloaded the results and performed further analysis by the R software. The criterion of function analysis was p-value cutoff < 0.05. Using metaspape, DEGs were studied for further (Fig 2).

Construction and analysis of PPI network

Online website STRING and cytoscape were used to build the PPI network and analyzed the functional pathways (Fig 3). Cytoscape functioned as a network graph, with differential expressed molecules represented as nodes and intermolecular interactions represented as links, that is, edges, between nodes [11]. In this study, the combined score represented intermolecular interactions and the protein pairs with combined score > 0.4 were selected to construct the PPI network [21]. Nodes were drawn in different sizes and colors, which represented the node degree and the regulation (up or down), respectively [18]. The hub genes were calculated by different methods with cytohubba [22]. The cytohubba is a plugin of cytoscape. It was widely used and essential for us to explore the most important node in diverse biological networks. CytoHubba included 11 topological analysis methods, while in this study we used Degree, Density of Maximum Neighborhood Component (DNMC), Maximal Clique Centrality (MCC) to identify candidate hub gene of prostate cancer.

Table 1. Characteristics of the microarray datasets.

Accession/ID	PMID	Experimental group	Control group
GSE3325	16286247	Prostate cancer tissue samples (7)	Benign prostate tissue samples (6)
GSE46602	26522007	Prostate cancer tissue samples (36)	Normal prostate gland samples (14)
GSE69223	26623558	Prostate cancer tissue samples (15)	Noncancerous tissue samples (15)
GSE104749	29285211	Prostate cancer tissue samples (4)	Benign prostate tissue samples (4)

<https://doi.org/10.1371/journal.pone.0247827.t001>

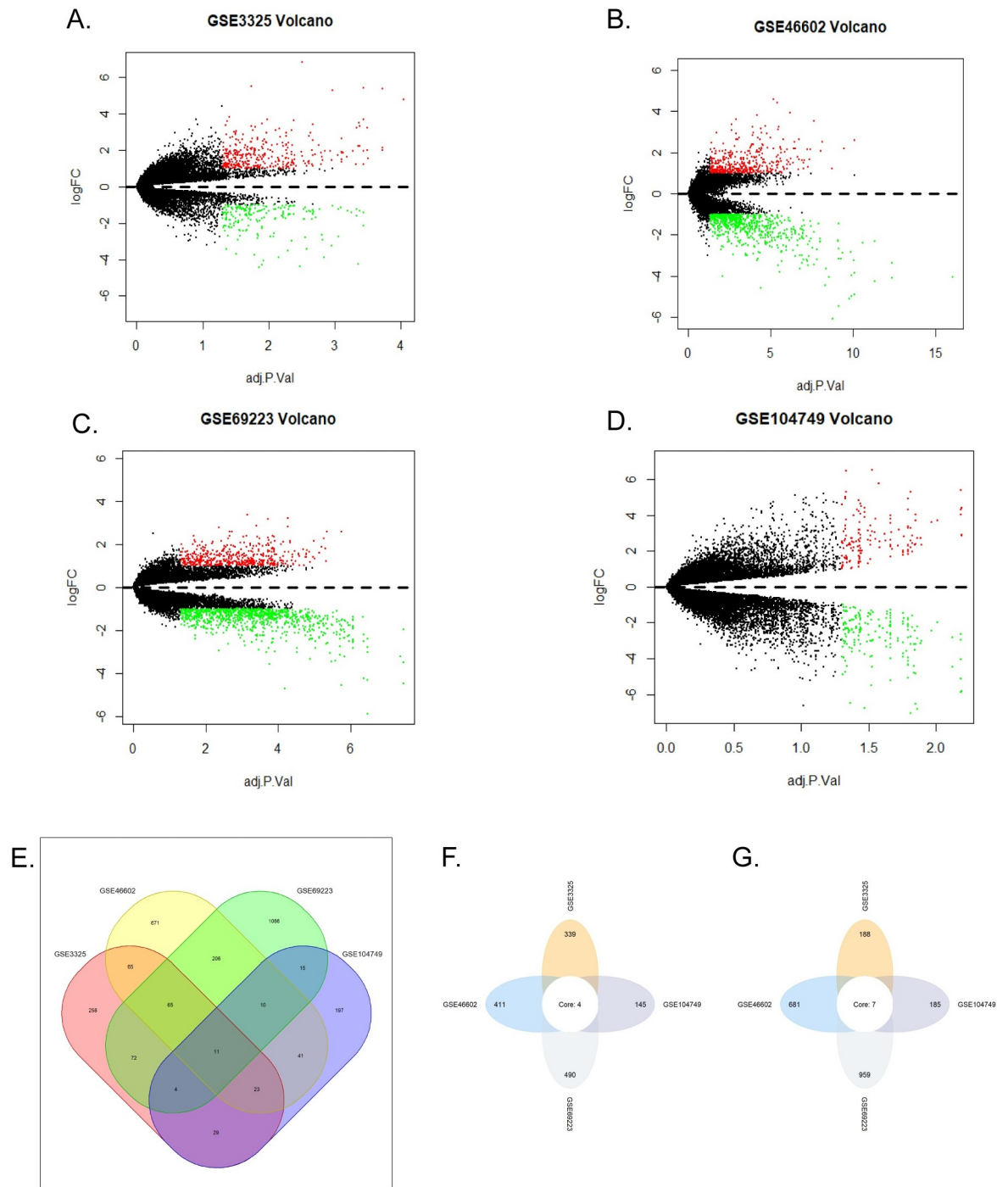


Fig 1. DEGs in GSE104749, GSE3325, GSE69223, GSE46602. (A-D) Volcano plots of all genes in four GEO databases. Red plots represented upregulated genes, green represented downregulated genes, the black plots represented the genes with no significant changes in expression. (A) GSE3325. (B) GSE46602. (C) GSE69223. (D) GSE104749. (E) Venn diagram of DEGs in four databases. The green, yellow, red, purple colors represented the DEGs of GSE46602, GSE69223, GSE104749, GSE3325, respectively. The number on the crossing area indicated the cross DEGs that were owned by different databases. (F) Upregulated genes co-expressed in the four databases. (G) Downregulated genes co-expressed in the four databases.

<https://doi.org/10.1371/journal.pone.0247827.g001>

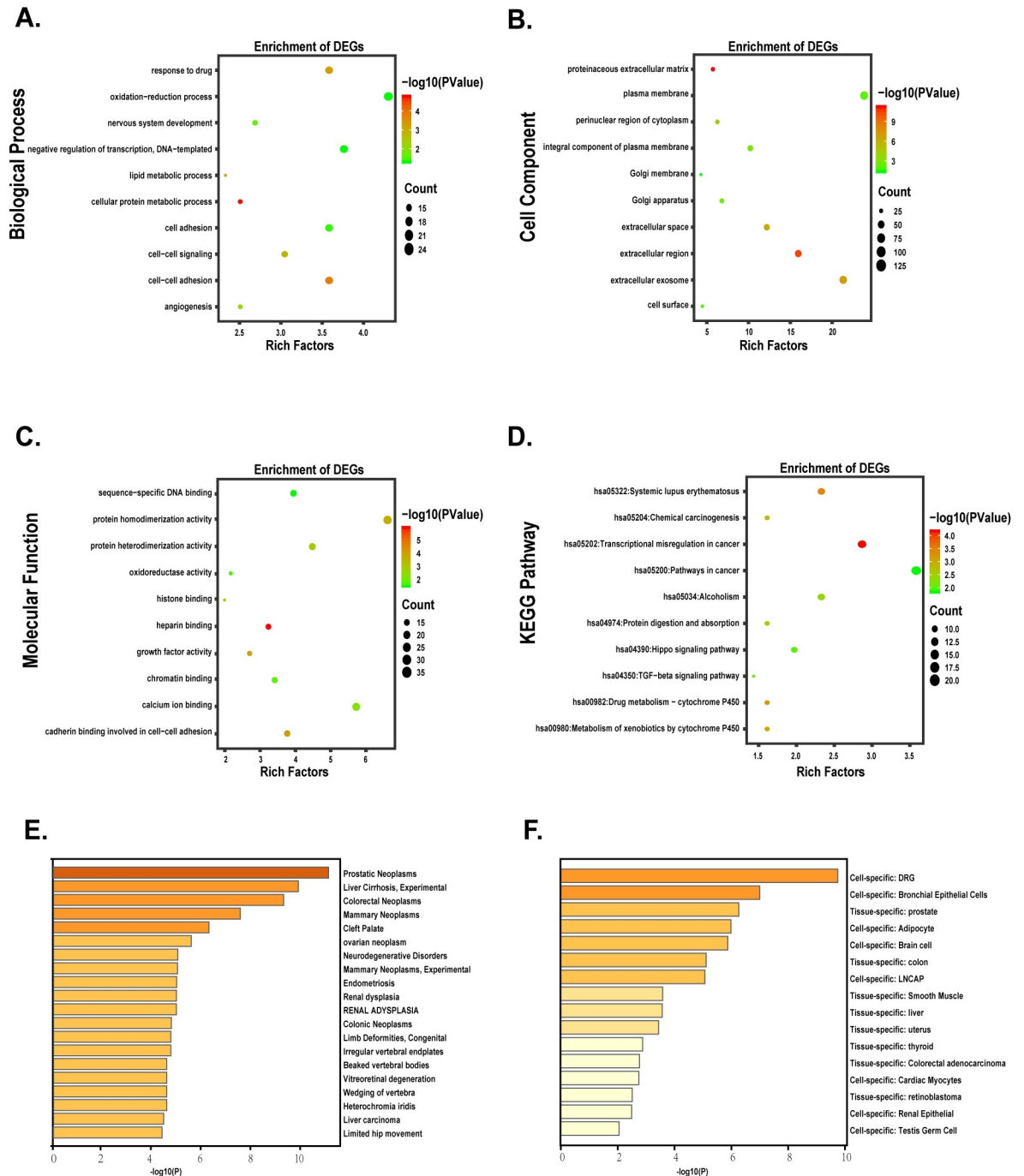


Fig 2. Functional analysis of DEGs. (A-D) The bubble plots indicated the results of function enrichment analysis in DEGs. (A) Biological process. (B) Cell component. (C) Molecular function. (D) KEGG pathways. (E-F) Different expression genes enriched in the specific tissues and cell lines.

<https://doi.org/10.1371/journal.pone.0247827.g002>

Survival analysis and expression validation of candidate hub genes

To further figure out the real hub gene from candidate hub genes of prostate cancer, Gene Expression Profiling Interactive Analysis (GEPIA, <http://gepia.cancer-pku.cn/>) [23], an online website based on The Cancer Genome Atlas [24] (TCGA) and the Genotype-Tissue

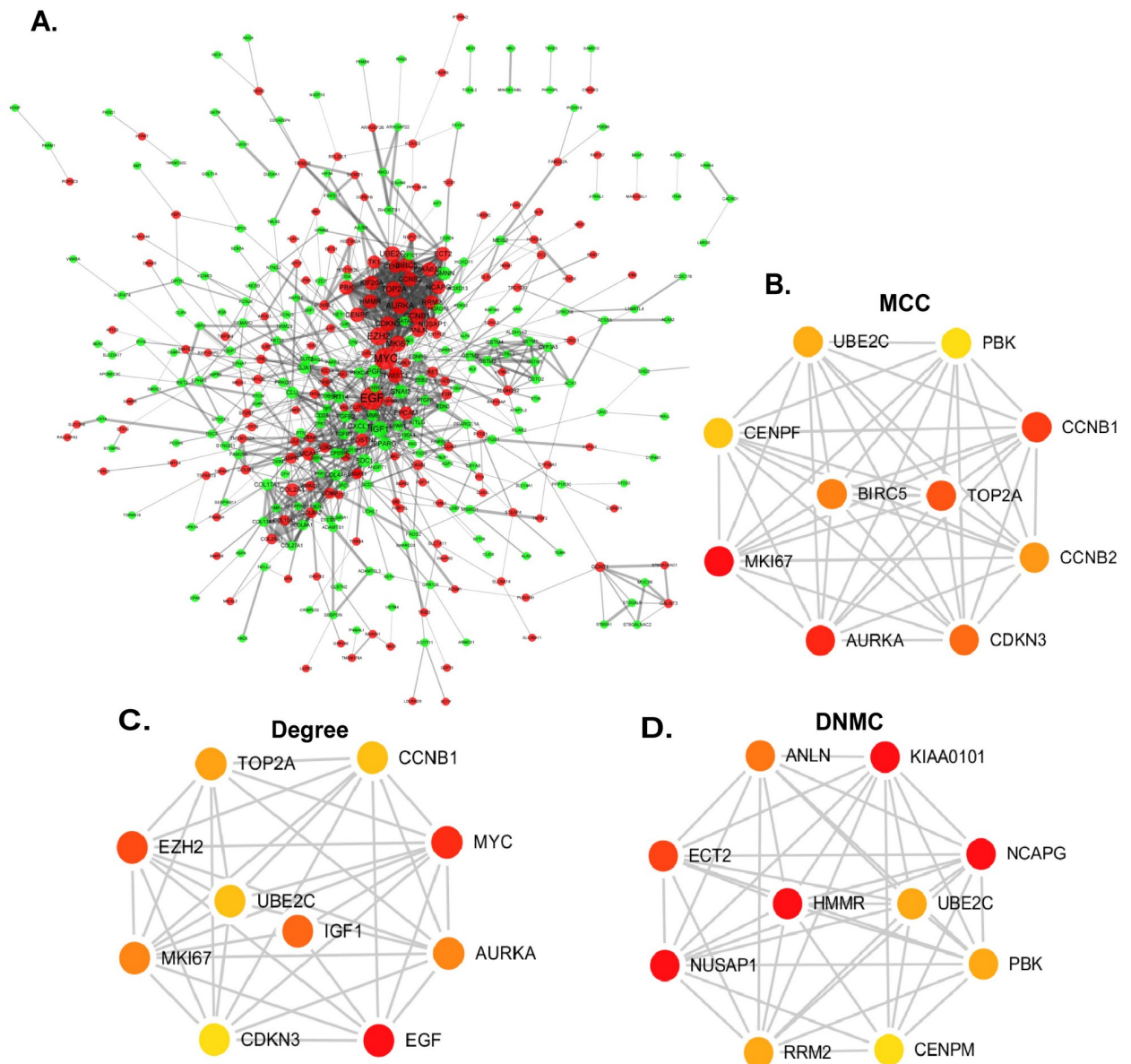


Fig 3. Construction of PPI network and identification of candidate hub genes. (A) PPI network. Red nodes represented upregulated genes while green represented downregulated genes. (B-D) Three different methods of cytoHubba were performed to identify the hub genes. The top ten nodes are shown with a color scheme from red (highly important) to green (important). (B) Top ten genes in the PPI network were calculated by MCC. (C) Top ten genes in the PPI network were selected by Degree as candidates. (D) Top ten genes in the PPI network were selected by DNMC.

<https://doi.org/10.1371/journal.pone.0247827.g003>

Expression (GTEx) databases [25], was not only performed to validate the expression of candidate hub genes between prostate cancer and benign prostate tissue, but also used to validate the correlations between disease-free survival and expression level of candidate hub genes (Fig 4).

Clinical parameters analysis

Considering the clinical application value of candidate hub genes, firstly, we analyzed the correlation between candidate hub genes and Gleason score of prostate cancer patients in GSE70770 (Fig 5). Secondly, we have validated the role of UBE2C in the T-stage and N-stage

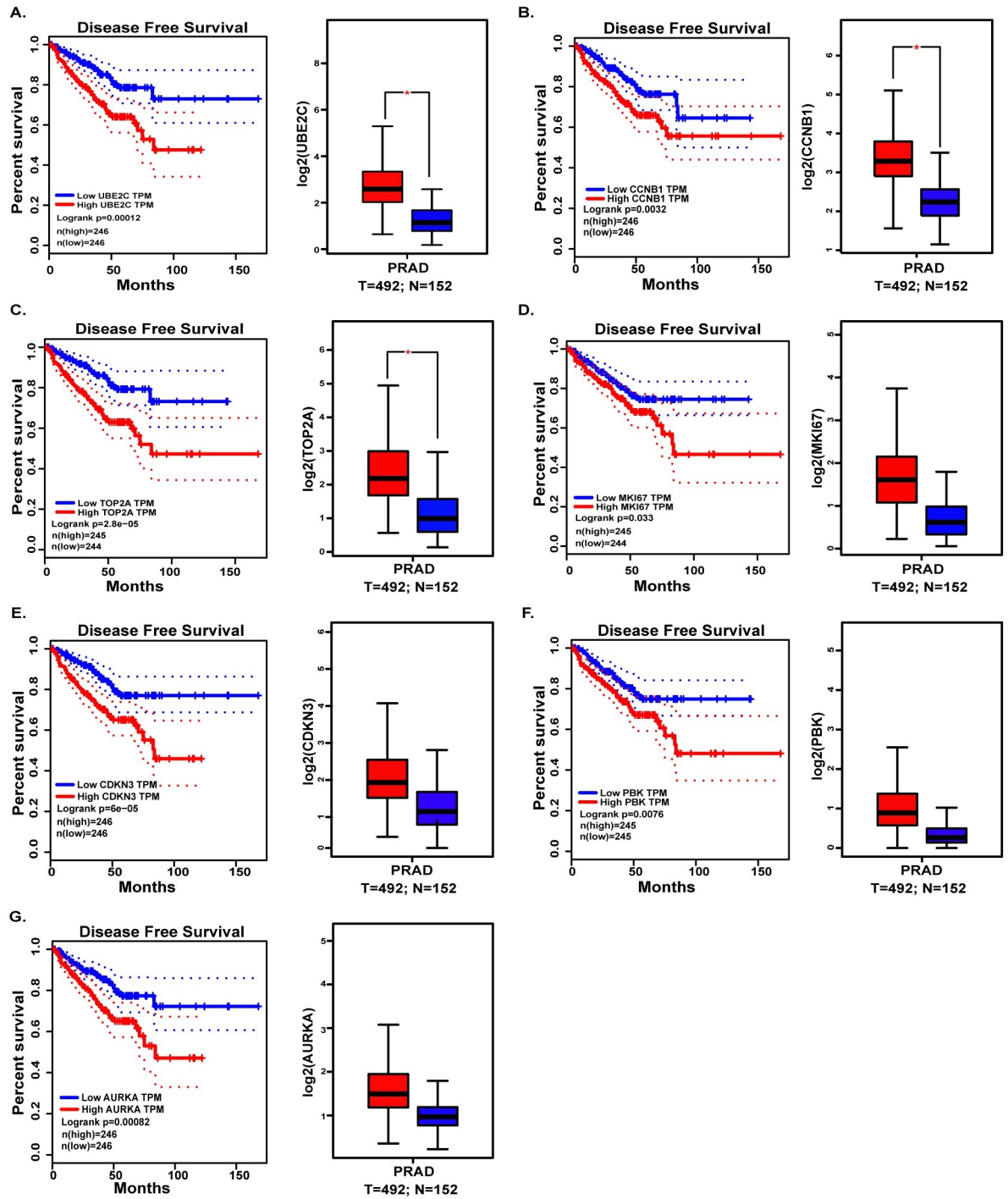


Fig 4. Validation the expression and disease-free survival (DFS) of candidate hub genes in TCGA. (A) UBE2C. (B) CCNB1. (C) TOP2A. (D) MKI67. (E) CDKN3. (F) PBK. (G) AURKA.

<https://doi.org/10.1371/journal.pone.0247827.g004>

of prostate cancer (Fig 6). Then we performed the disease-free survival and expression analysis of UBE2C in GSE70770 (three samples was excluded because of the absence of the follow up months), GSE116918 and TCGA by KM curve and boxplot (Figs 6 and 7). The correlations between clinical parameters and the expression of UBE2C included in Table 2 was done by

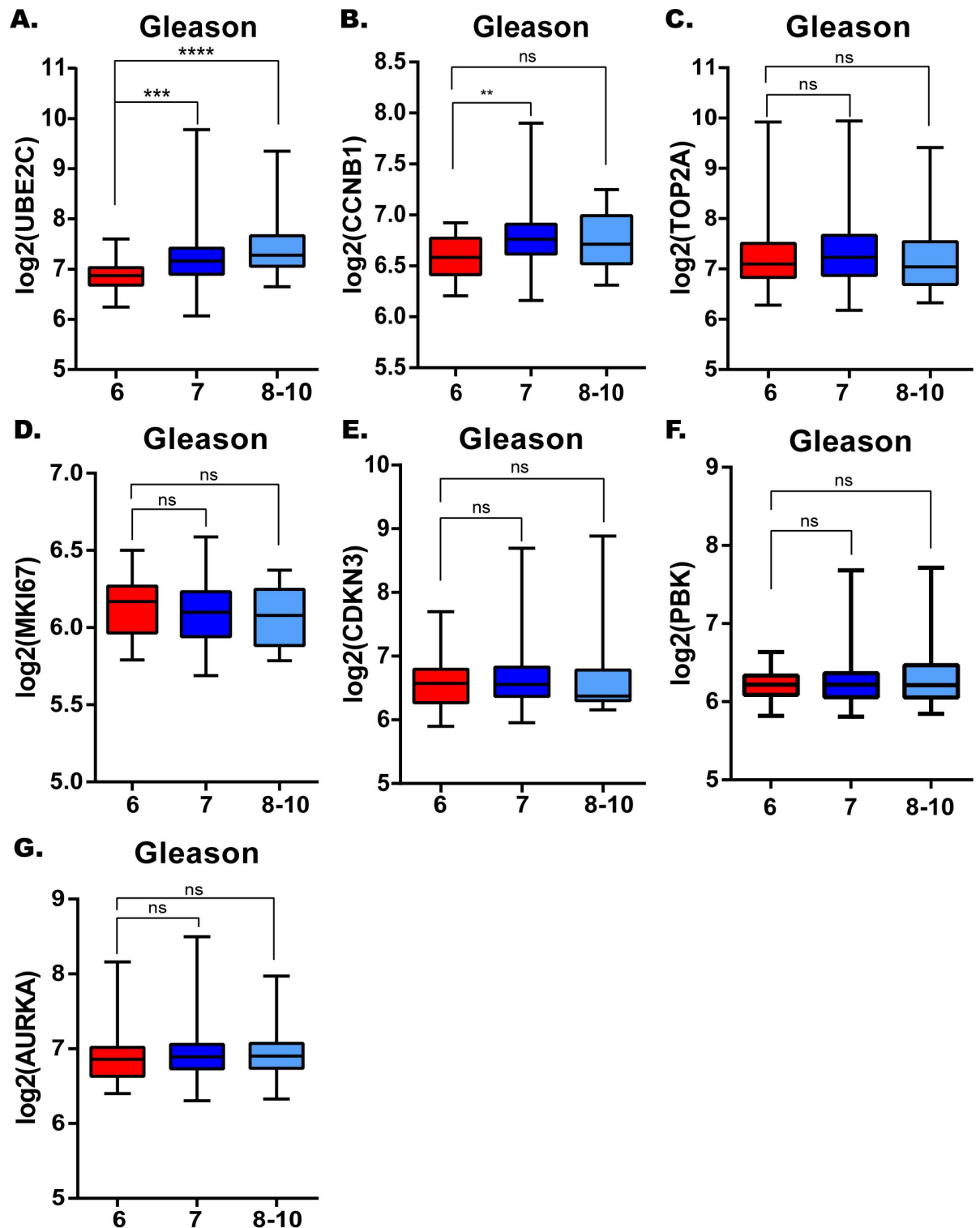


Fig 5. The correlations between candidate hub genes and Gleason score. UBE2C. (B) CCNB1. (C) TOP2A. (D) MKI67. (E) CDKN3. (F) PBK. (G) AURKA. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$.

<https://doi.org/10.1371/journal.pone.0247827.g005>

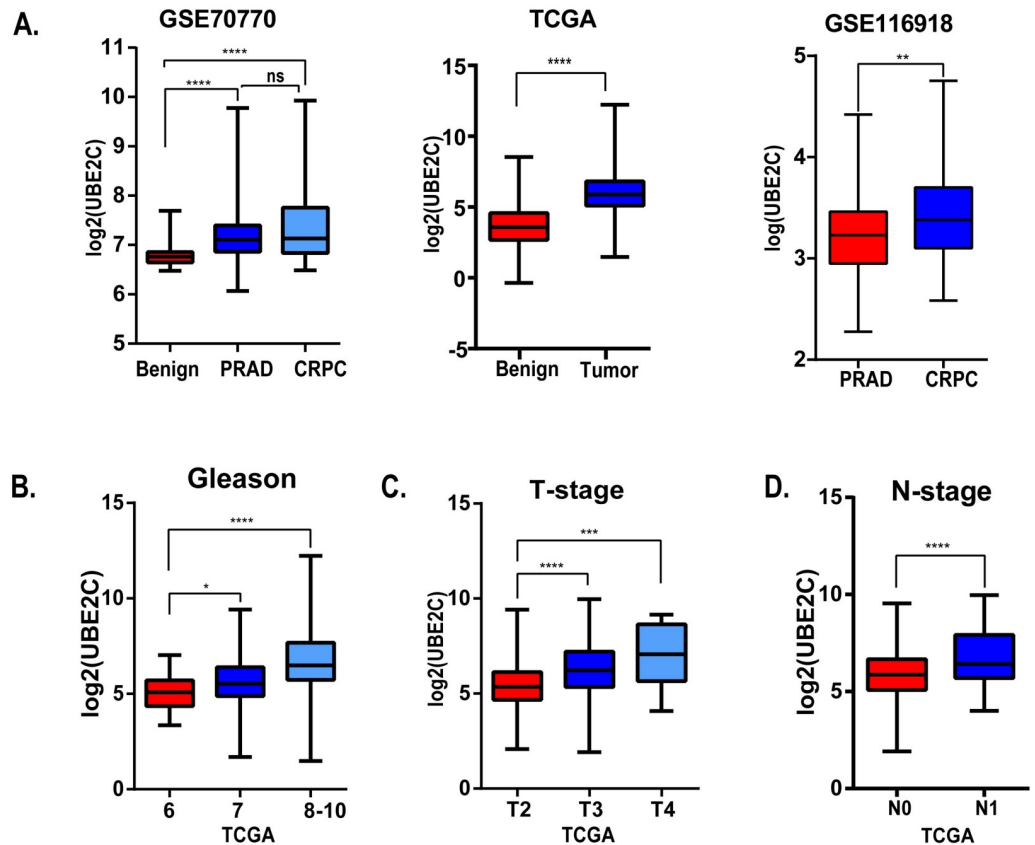


Fig 6. The correlations between UBE2C and clinical parameters. (A) The expression level of UBE2C of benign prostate tissue, prostate cancer and castration-resistant prostate cancer in GSE70770 (left), TCGA (middle) and GSE116918 (right). (B) The relationship between Gleason score and UBE2C in TCGA. (C) The relationship between T-stage and UBE2C in TCGA. (D) The relationship between N-stage and UBE2C in TCGA. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$.

<https://doi.org/10.1371/journal.pone.0247827.g006>

SPSS version 19.0. Continuous variables were analyzed by independent t tests, as for categorical variables which were analyzed by chi-square test. By performing univariate and multivariate cox regression analysis (Table 3), the correlations between clinical parameters and disease-free survival were done by R software. Rms package of R software was used to do nomogram plots analysis (S2 Fig). ROC curve analysis was done with survivalROC and pROC package (Figs 7 and S3).

GSEA

The GEO dataset (GSE70770) contained 206 primary prostate cancers, which were divided into two groups based on the median expression of UBE2C, high versus low. To figure out the function and effect of UBE2C, GSEA was used to do the enrichment analysis. The criteria cut-off set up as P-value < 0.05 and false discovery rate (FDR) < 0.25 (Fig 8).

Results

DEGs in prostate cancer

As shown in volcano plots (Fig 1A–1D), we got 527 DEGs in GSE3325, 1092 DEGs in GSE46602, 330 DEGs in GSE104749 and 1449 DEGs in GSE69223. As shown in Venn diagram

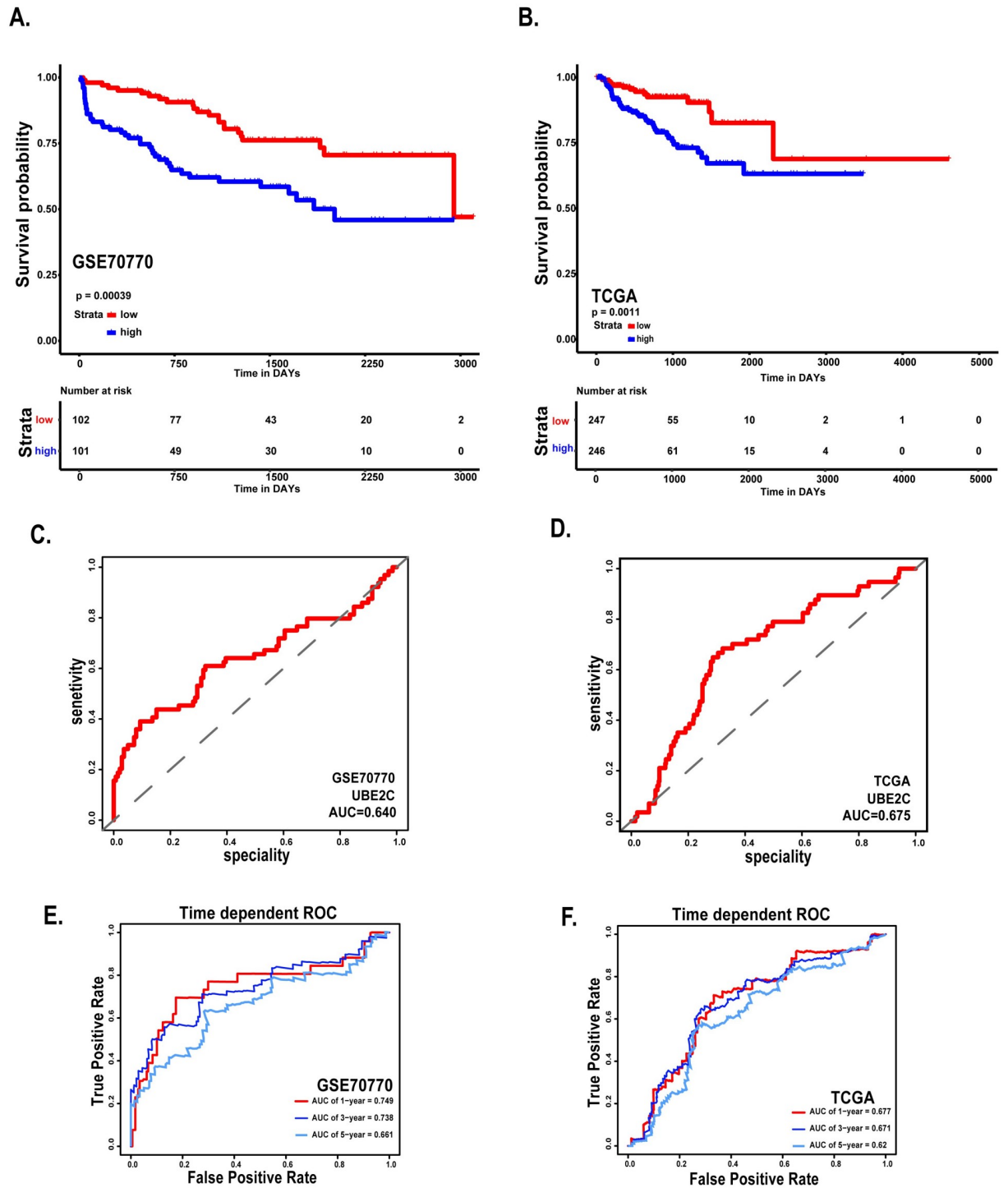


Fig 7. Survival analysis and ROC curve of UBE2C in GSE70770 and TCGA. (A) Disease free survival of patients based on the expression of UBE2C in GSE70770. (B) Disease free survival of patients based on the expression of UBE2C in TCGA. (C) ROC curve of UBE2C in GSE70770. (D) ROC curve of UBE2C in TCGA. (E-F) Time-dependent ROC curves. The AUCs at 1, 3, and 5 years were used to assess prognostic accuracy (left: GSE70770, right: TCGA). *: p<0.05; **: p<0.01; ***: p<0.001; ****: p<0.0001.

<https://doi.org/10.1371/journal.pone.0247827.g007>

Table 2. The relationships between UBE2C and clinical parameters (n = 203).

Clinical parameters		Low UBE2C (n = 102)	High UBE2C (n = 101)	P value
Age	Mean±Std. Deviation	60.31±6.567	60.59±6.685	0.863
	Median(Minimum ~Maximum)	62(41~72)	61(42~73)	
Gleason Score	≤6	31(30.4%)	6(5.9%)	<0.0001
	7 = 3+4	46(45.1%)	55(54.5%)	
	7 = 4+3	16(15.7%)	23(22.8%)	
	8–10	7(6.9%)	17(16.8%)	
	Unknown	2(2.0%)	0(0.0%)	
PSA	Mean±Std. Deviation	9.8117±12.04088	9.2606±5.34981	0.187
	<4	5(4.9%)	4(4.0%)	0.798
	4–10	70(68.6%)	66(65.3%)	
	>10	25(24.5%)	30(29.7%)	
Extra-capsular extension (ECE)	Yes	48(47.1%)	71(70.3%)	<0.001
	No/Unknown	54(52.9%)	30(29.7%)	
Positive surgical margins (PSM)	Yes	30(29.4%)	63(62.4%)	0.215
	No/Unknown	72(70.6%)	38(37.6%)	
Biochemical relapse (BCR)	Yes	23(22.55%)	41(40.6%)	0.006
	No/Unknown	79(77.5%)	60(59.4%)	

<https://doi.org/10.1371/journal.pone.0247827.t002>

(Fig 1E), a total of 243 upregulated genes and 298 downregulated genes were identified. In addition, transient receptor potential melastatin-4 (TRPM4), distal-less homeobox 1 (DLX1), collagen triple helix repeat containing 1 (CTHRC1) and asporin (ASPN) are high expression in four datasets, while K+ channel gene (KCN3), Alcohol oxidase I (AOX1), HUT11/UT-B/JK (SLC14A1), collagen type XIII alpha 1 (COL13A1), transcription elongation factor A (SII)-like 2 (TCEAL2), homeobox D10 (HOXD10), carboxypeptidase A6 (CPA6) are decreased in all datasets. (Fig 1F and 1G).

GO and KEGG pathway analysis

To figure out the function of these DEGs, as described in S1 Table, DAVID was used to perform GO and KEGG pathway and the results were presented in S2 Table. As shown in Fig 2,

Table 3. Univariate and multivariate Cox regression analysis for prostate cancer relapse-free survival.

Variate		Univariate	P value	multivariate	P value
		HR (95%CI)		HR (95%CI)	
Age		1.011(0.942–1.084)	0.763		
UBE2C		3.142(2.218–4.449)	0.0001	2.796(1.762–4.436)	0.0001
PSA		1.022(1.005–1.038)	0.01	1.014(0.992–1.037)	0.217
Gleason	≤ 6	1		1	
	7 = 3+4	6.186(1.462–26.175)	0.013	3.576(0.811–15.770)	0.092
	7 = 4+3	16.085(3.752–68.947)	<0.0001	8.974(1.984–40.601)	0.004
	8–10	32.987(7.539–144.331)	<0.0001	12.568(2.602–60.703)	0.002
ECE	YES	1		1	
	NO/unknown	0.393(0.227–0.683)	0.001	0.619(0.335–1.144)	0.126
PSM	YES	1		1	
	NO/unknown	0.455(0.278–0.744)	0.002	0.499(0.295–0.842)	0.009

ECE: Extra-capsular extension PSM: Positive surgical margins HR: Hazard ratio CI: Confidence interval.

<https://doi.org/10.1371/journal.pone.0247827.t003>

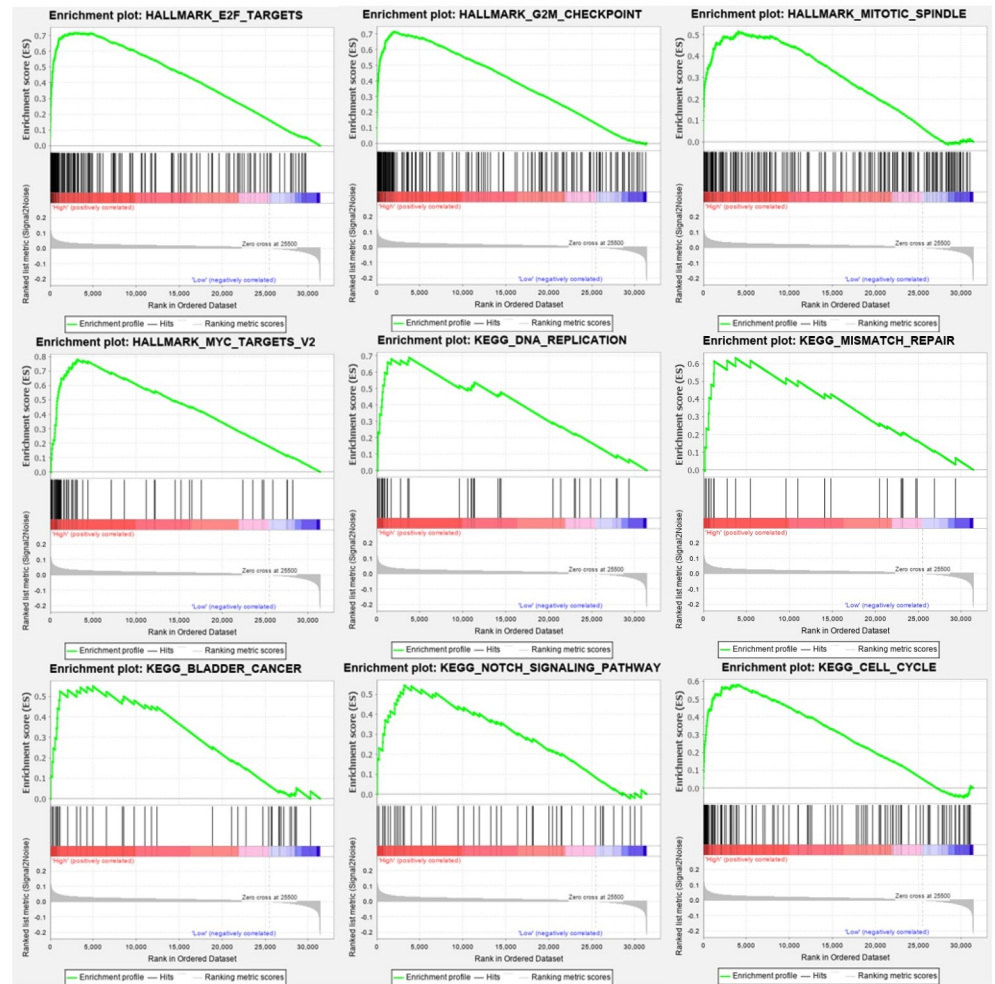


Fig 8. Genes enrichment analysis. The results above all meet the standards ($FDR < 0.25$, $P\text{-value} < 0.05$). The results showed that high expression of UBE2C were enriched in G2M check point, cell cycle, DNA replication, E2F targets, MYC targets, Notch signaling pathway.

<https://doi.org/10.1371/journal.pone.0247827.g008>

biological process analysis indicated DEGs were significantly enriched in the oxidation-reduction process, negative regulation of transcription, cell-cell signaling, lipid metabolic process, cell adhesion. For the cellular component, the most significantly altered pathway was proteinaeous extracellular matrix, extracellular region. As for molecular function, cadherin binding involved in cell-cell adhesion, oxidoreductase activity was significantly enriched. KEGG pathway analysis showed that transcriptional dysregulation in cancer, TGF-beta signaling pathway, Hippo signaling pathway were essential for prostate cancer. With metascape, the results indicated that DEGs were obviously correlated with the development of prostate neoplasms (Fig 2E and 2F).

PPI network construction and hub genes validation

In order to screen the hub genes, the different expression genes were uploaded to STRING, then PPI network was further constructed by Cytoscape (Fig 3). In order to identify the hub genes of prostate cancer, DEGs were calculated by Maximal Clique Centrality (MCC) [26], Degree method, Density of Maximum Neighborhood Component (DNMC) [22] to screen the

top ten genes. The 7 genes, UBE2C, PBK, CCNB1, CDKN3, TOP2A, AURKA and MKI67, presenting in at least two methods, were recognized as candidate hub genes. But only UBE2C, screening in all methods, was recognized as of higher worth. In order to verify the role and function of 7 genes in the progression and prognosis of prostate cancer, expression level and survival analysis was performed by GEPIA (Fig 4). The results showed that 7 genes were all associated with prostate cancer prognosis, but only UBE2C, TOP2A and CCNB1 were high expression in prostate cancer samples than normal prostate gland samples, the expression of PBK, CDKN3, AURKA, MKI67 were not.

Clinical information of the hub gene

Considering the clinical value of 7 genes, we analyzed the correlation between candidate hub genes and Gleason score of prostate cancer. As the results showed that, candidate hub genes, in particular, UBE2C was positive related with Gleason score in GSE70770 (Figs 5 and 6B). The higher expression level of UBE2C indicated the worse Gleason score. Moreover, the expression level of UBE2C increased as T-stage and N-stage of prostate cancer improved (Fig 6C and 6D). According to the validation results, UBE2C, which was highly expressed in prostate cancer (Fig 6A), in particular castration-resistant prostate cancer. And the results showed that UBE2C was essential for predicting the disease-free survival of prostate cancer and may play an important role in prostate cancer (Figs 7 and S4). UBE2C was confirmed as the hub gene of prostate cancer. As shown in Table 2, based on the median expression of UBE2C, primary prostate cancer samples in GSE70770 were divided into low versus high group. As we expected, extra-capsular extension (ECE), Gleason score and biochemical relapse (BCR) had remarkable differences between the two groups, while no significant differences were observed in age, PSA and positive surgical margins (PSM). To figure out whether UBE2C could function as the independent prognostic factor of prostate cancer, the effect of clinical phenotype on relapse-free survival was presented by Cox regression (Table 3). According to the results, UBE2C functioned as a risk factor in prostate cancer progression (HR 2.796; 95%CI 1.762–4.436; $P < 0.0001$). ROC curve was applied to present the predictive accuracy of UBE2C (AUC = 0.64 in GSE70770, AUC = 0.675 in TCGA) in DFS analysis (Fig 7). In GSE70770, the AUC of prostate cancer BCR at 1-year is 0.749, 0.738 for 3-year, and 0.661 for 5-year. And the results in TCGA (1-year AUC: 0.677, 3-year AUC: 0.671, 5-year AUC: 0.62) were similar to GSE70770. UBE2C combined with clinical features achieved a more remarkable area under the curve of ROC which was significantly superior to UBE2C alone (S3 Fig). Using nomogram that integrated clinicopathological features (S5 Fig), we predict survival probabilities of patients to improve the patients' living condition. For localized prostate cancer, radical prostatectomy combined with radiation has become the first choice, although it can bring some adverse effects to patients. Previous study has showed UBE2C was highly correlated with chemoresistance and radiotherapy resistance of prostate cancer [27–30]. According to our results, UBE2C highly expressed in prostate cancer, especially in castration-resistant prostate cancer and UBE2C was correlated with the neuroendocrine prostate cancer biomarkers, such as RB1 and LDHA (S4 Fig). It indicated that targeting UBE2C may provide new therapy idea for prostate cancer, especially for castration-resistant prostate cancer.

GSEA

As shown in Fig 8, high expression of UBE2C were positive with G2M check point, cell cycle, DNA replication. Furthermore, E2F targets, MYC targets, Notch signaling pathway were also significantly enriched accompanying with UBE2C expression. As we expected, UBE2C is

positively correlated with bladder cancer progression as well as renal cell carcinoma, pancreatic cancer.

Discussion

Ubiquitin, which was a conserved protein containing 76 amino acids, could be included in the process of post-translation for protein degradation [31]. We confirmed that ubiquitin could be conjugated to lysine residue of target protein through multi-step reactions of enzymes [31–33]. Ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2) and ubiquitin ligases (E3) were all involved in the ubiquitination. As we all know, ubiquitination, which was one of the post-translational modifications, began with activation of ubiquitin molecule and was terminated by the linkage of ubiquitin to the target protein [32,33]. Ubiquitination was essential for the regulation of a number of cellular processes among almost all mammalian cells [34,35]. The modification involved in the progression of many severe diseases, like infection, inflammation response, cancer, neurodegeneration [34–36]. It provided us with a novel strategy for the therapeutic intervention of tumor by regulating the activity of ubiquitin enzyme [35].

Based on current study, UBE2C, which could not only drive prostate neoplasms progression but also could predict the prognosis of prostate cancer, was identified as the hub gene of prostate cancer. We have confirmed that the level of UBE2C was paralleled with the Gleason score of prostate cancer, early biochemical recurrence and poor clinical outcomes. As the results showed, the higher UBE2C level was, the worse Gleason score would be. UBE2C, which belongs to the family of E2 ubiquitin-conjugating enzyme, has played a crucial role in inducing protein degradation through ubiquitin-proteasome proteolytic (UPP) pathway in cooperation with anaphase promoting complex (APC) [37,38]. UBE2C was highly expressed in diverse tumors when compared with respective normal tissue [38,39], such as breast cancer, lung cancer, colon cancer, liver cancer, thyroid cancer, prostate cancer [40]. The mechanisms of UBE2C in promoting and regulating the development of prostate cancer deserve to be explored. And there have been several researches focused on the mechanism of UBE2C in prostate cancer. Other research has confirmed the level of UBE2C, which could be regulated by miR-381-3p, was positive correlated with the proliferation of prostate cancer cells [39]. Post-translational modification of Mediator 1 (MED1) T1032 phosphorylation, which was regulated by PI3K/AKT signaling pathway [41], promoted the expression of UBE2C in prostate cancer and enhanced the role of UBE2C in promoting the proliferation of prostate cancer and in driving prostate cancer progression. Conspicuously, in PC-3 cells, UBE2C could recruit FOXA1 to its enhancers, that was mostly likely correlated with the expression of UBE2C [41] and promoted the progression of castration-resistant prostate cancer. Other research has reported that FOXM1, E2F1 and RAD51 could bind to the enhancer regions and promoter regions of UBE2C in order to regulate the expression of UBE2C [40]. Apparently, the level of UBE2C showed strong relationship with the differentiation and progression of prostate cancer and conspicuously correlated with the prognosis. Compared with low risk androgen-dependent prostate carcinoma (ADPC), UBE2C prominently upregulated in the fatal castration-resistant prostate cancer (CRPC) [37]. ADT has been considered as the first-line treatment for advanced prostate cancer. Although in a few months later, prostate cancer would progress to CRPC which had no response to the ADT [42]. The pathogenesis of CRPC was most closely to AR mutation, AR amplification, the presence of AR-V7 and abnormal activation of androgen receptors downstream signals [43]. Compared with full-length AR (AR-FL), the level of UBE2C was accurately correlated with the expression of AR variant 7 (AR-V7), which could regulate the expression of UBE2C [42]. It revealed that UBE2C was the downstream target of

AR-V7, which would give us a brand-new way to explore the therapy for CRPC by targeting the activity of UBE2C.

Through comprehensive bioinformatics analysis, we concluded that UBE2C could drive the progression of prostate cancer, but the analysis in this paper still had some certain limitations. First, the role of UBE2C in prostate cancer has not been proven in vivo or in vitro. It is essential for us to verify the importance of UBE2C in prostate cancer and explore the mechanism of UBE2C in regulating the development of prostate cancer by experimental approach. Second, the study of chemotherapy and radiotherapy resistance of prostate cancer patients is very important. It is of great worth to study the relationship between the expression of UBE2C and the treatment resistance of prostate cancer. So we should pay attention to performing clinical trials and we also should focus on the prevention, the treatment and the care of cancer.

Supporting information

S1 Fig. The flowchart of inclusion criteria of GEO gene expression profiles.

(DOCX)

S2 Fig. DEGs in GSE104749, GSE3325, GSE69223, GSE46602 and pathway analysis. (A-D) Heatmap plots of four GEO databases. Red plots symbolized upregulation genes, green represented downregulation genes, the black plots represented the genes with no significant expression change. (A) GSE3325 (B) GSE69223 (C) GSE104749 (D) GSE46602. (E) The barplot of GO and KEGG pathway analysis of DEGs.

(DOCX)

S3 Fig. The ROC curve of UBE2C combined with clinicopathological phenotype in GSE70770. (A) UBE2C and Gleason score (AUC = 0.7475). (B) UBE2C, Gleason score and age (AUC = 0.8202). (C) UBE2C, Gleason score, age and T-stage (AUC = 0.8459). (D) UBE2C and MKI67 (AUC = 0.726) (E) UBE2C and CDKN3 (AUC = 0.677). (F) UBE2C and CCNB1 (AUC = 0.674). (G) UBE2C and TOP2A (AUC = 0.649). (H) UBE2C and PBK (AUC = 0.645). (I). UBE2C and AURKA (AUC = 0.659).

(DOCX)

S4 Fig. The role of UBE2C in castration-resistant prostate cancer. (A) DFS in GSE116918. (B) ROC curve of UBE2C in GSE116918. (C) Time dependent ROC of UBE2C in GSE116918. (D) The correlation of UBE2C and RB1 in castration-resistant prostate cancer. (E) The correlation of UBE2C and LDHA in castration-resistant prostate cancer.

(DOCX)

S5 Fig. Nomogram of prostate cancer to predict 3-year, 5-year survival probability.

(A) Nomogram of prostate cancer. B-C: Nomogram's calibration chart. The dotted line represents an ideal reference line where the predicted probability matches the proportion of observations. B: Nomogram-Predicted Probability of 3-year survival. C: Nomogram-Predicted Probability of 5-year survival.

(DOCX)

S1 Table. Identification of DEGs in prostate cancer among four databases.

(DOCX)

S2 Table. GO and KEGG signaling pathways (TOP 10).

(DOCX)

Acknowledgments

We appreciated the public database and online tools performed in this research such as GEO, DAVID, Metascape, Cytoscape, TCGA, GSEA.

Author Contributions

Conceptualization: Guoping Ren.

Formal analysis: Yan Wang.

Methodology: Yan Wang, Guoping Ren.

Validation: Yan Wang, Qiusu Tang.

Visualization: Jili Wang, Qiusu Tang, Guoping Ren.

Writing – original draft: Yan Wang.

Writing – review & editing: Yan Wang, Jili Wang, Guoping Ren.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* 2019; 69(1):7–34. <https://doi.org/10.3322/caac.21551> PMID: 30620402
2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin.* 2016; 66(2):115–32. <https://doi.org/10.3322/caac.21338> PMID: 26808342
3. Litwin MS, Tan HJ. The Diagnosis and Treatment of Prostate Cancer: A Review. *JAMA.* 2017; 317(24):2532–42. <https://doi.org/10.1001/jama.2017.7248> PMID: 28655021
4. Chang AJ, Autio KA, Roach M, Scher HI. High-risk prostate cancer-classification and therapy. *Nat Rev Clin Oncol.* 2014; 11(6):308–23. <https://doi.org/10.1038/nrclinonc.2014.68> PMID: 24840073
5. Intasqui P, Bertolla RP, Sadi MV. Prostate cancer proteomics: clinically useful protein biomarkers and future perspectives. *Expert Rev Proteomic.* 2018; 15(1):65–79. <https://doi.org/10.1080/14789450.2018.1417846> PMID: 29251021
6. Fang EH, Zhang XQ, Wang Q, Wang DM. Identification of prostate cancer hub genes and therapeutic agents using bioinformatics approach. *Cancer Biomark.* 2017; 20(4):553–61. <https://doi.org/10.3233/CBM-170362> PMID: 28800317
7. Barros-Silva JD, Linn DE, Steiner I, Guo GJ, Ali A, Pakula H, et al. Single-Cell Analysis Identifies LY6D as a Marker Linking Castration-Resistant Prostate Luminal Cells to Prostate Progenitors and Cancer. *Cell Rep.* 2018; 25(12):3504–+. <https://doi.org/10.1016/j.celrep.2018.11.069> PMID: 30566873
8. Dennis G Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 2003; 4(5):P3. PMID: 12734009
9. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 2017; 45(D1):D362–D8. <https://doi.org/10.1093/nar/gkw937> PMID: 27924014
10. Zhou YY, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun.* 2019; 10. <https://doi.org/10.1038/s41467-019-09234-6> PMID: 30944313
11. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003; 13(11):2498–504. <https://doi.org/10.1101/gr.1239303> PMID: 14597658
12. Otasek D, Morris JH, Boucas J, Pico AR, Demchak B. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biol.* 2019; 20(1). <https://doi.org/10.1186/s13059-019-1758-4> PMID: 31477170
13. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc.* 2019; 14(2):482–517. <https://doi.org/10.1038/s41596-018-0103-9> PMID: 30664679
14. Lu WZ, Ding Z. Identification of key genes in prostate cancer gene expression profile by bioinformatics. *Andrologia.* 2019; 51(1). <https://doi.org/10.1111/and.13169> PMID: 30311263

15. Song ZS, Huang Y, Zhao Y, Ruan HL, Yang HM, Cao Q, et al. The Identification of Potential Biomarkers and Biological Pathways in Prostate Cancer. *J Cancer*. 2019; 10(6):1398–408. <https://doi.org/10.7150/jca.29571> PMID: 31031850
16. Liu S, Wang WX, Zhao Y, Liang KG, Huang YJ. Identification of Potential Key Genes for Pathogenesis and Prognosis in Prostate Cancer by Integrated Analysis of Gene Expression Profiles and the Cancer Genome Atlas. *Front Oncol*. 2020; 10. <https://doi.org/10.3389/fonc.2020.00809> PMID: 32547947
17. Foj L, Filella X. Identification of Potential miRNAs Biomarkers for High-Grade Prostate Cancer by Integrated Bioinformatics Analysis. *Pathol Oncol Res*. 2019; 25(4):1445–56. <https://doi.org/10.1007/s12253-018-0508-3> PMID: 30367364
18. Wang JL, Wang Y, Ren GP. Identification of PTPRR and JAG1 as key genes in castration-resistant prostate cancer by integrated bioinformatics methods. *J Zhejiang Univ-Sc B*. 2020; 21(3):246–55. <https://doi.org/10.1631/jzus.B1900329> PMID: 32133801
19. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, et al. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res*. 2004; 32:D258–D61. <https://doi.org/10.1093/nar/gkh036> PMID: 14681407
20. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000; 28(1):27–30. <https://doi.org/10.1093/nar/28.1.27> PMID: 10592173
21. Wang YT, Wang JF, Yan KX, Lin JX, Zheng ZH, Bi JB. Identification of core genes associated with prostate cancer progression and outcome via bioinformatics analysis in multiple databases. *Peerj*. 2020; 8. <https://doi.org/10.7717/peerj.8786> PMID: 32266115
22. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *Bmc Syst Biol*. 2014; 8. <https://doi.org/10.1186/1752-0509-8-S4-S11> PMID: 25521941
23. Tang ZF, Li CW, Kang BX, Gao G, Li C, Zhang ZM. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017; 45(W1):W98–W102. <https://doi.org/10.1093/nar/gkx247> PMID: 28407145
24. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013; 45(10):1113–20. <https://doi.org/10.1038/ng.2764> PMID: 24071849
25. Ardlie KG, DeLuca DS, Segre AV, Sullivan TJ, Young TR, Gelfand ET, et al. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*. 2015; 348(6235):648–60. <https://doi.org/10.1126/science.1262110> PMID: 25954001
26. Li CZ, Xu JC. Feature selection with the Fisher score followed by the Maximal Clique Centrality algorithm can accurately identify the hub genes of hepatocellular carcinoma. *Sci Rep-Uk*. 2019; 9. <https://doi.org/10.1038/s41598-019-53471-0> PMID: 31754223
27. Planche A, Bacac M, Provero P, Fusco C, Delorenzi M, Stehle JC, et al. Identification of Prognostic Molecular Features in the Reactive Stroma of Human Breast and Prostate Cancer. *Plos One*. 2011; 6(5). <https://doi.org/10.1371/journal.pone.0018640> PMID: 21611158
28. Lee CH, Ku JY, Ha JM, Bae SS, Lee JZ, Kim CS, et al. Transcript Levels of Androgen Receptor Variant 7 and Ubiquitin-Conjugating Enzyme 2C in Hormone Sensitive Prostate Cancer and Castration-Resistant Prostate Cancer. *Prostate*. 2017; 77(1):60–71. <https://doi.org/10.1002/pros.23248> PMID: 27550197
29. Gu P, Chen X, Xie RH, Han JL, Xie WB, Wang B, et al. lncRNA HOXD-AS1 Regulates Proliferation and Chemo-Resistance of Castration-Resistant Prostate Cancer via Recruiting WDR5. *Mol Ther*. 2017; 25(8):1959–73. <https://doi.org/10.1016/j.ymthe.2017.04.016> PMID: 28487115
30. Tzelepi V, Zhang JX, Lu JF, Kleb B, Wu GL, Wan XH, et al. Modeling a Lethal Prostate Cancer Variant with Small-Cell Carcinoma Features. *Clin Cancer Res*. 2012; 18(3):666–77. <https://doi.org/10.1158/1078-0432.CCR-11-1867> PMID: 22156612
31. Popovic D, Vucic D, Dikic I. Ubiquitination in disease pathogenesis and treatment. *Nat Med*. 2014; 20(11):1242–53. <https://doi.org/10.1038/nm.3739> PMID: 25375928
32. Grumati P, Dikic I. Ubiquitin signaling and autophagy. *J Biol Chem*. 2018; 293(15):5404–13. <https://doi.org/10.1074/jbc.TM117.000117> PMID: 29187595
33. Swatek KN, Komander D. Ubiquitin modifications. *Cell Res*. 2016; 26(4):399–422. <https://doi.org/10.1038/cr.2016.39> PMID: 27012465
34. van Wijk SJ, Fulda S, Dikic I, Heilemann M. Visualizing ubiquitination in mammalian cells. *Embo Rep*. 2019; 20(2). <https://doi.org/10.15252/embr.201846520> PMID: 30665942
35. Rape M. Ubiquitylation at the crossroads of development and disease. *Nat Rev Mol Cell Bio*. 2018; 19(1):59–70. <https://doi.org/10.1038/nrm.2017.83> PMID: 28928488

36. Song L, Luo ZQ. Post-translational regulation of ubiquitin signaling. *J Cell Biol.* 2019; 218(6):1776–86. <https://doi.org/10.1083/jcb.201902074> PMID: 31000580
37. Mo CH, Gao L, Zhu XF, Wei KL, Zeng JJ, Chen G, et al. The clinicopathological significance of UBE2C in breast cancer: a study based on immunohistochemistry, microarray and RNA-sequencing data. *Cancer Cell Int.* 2017; 17:83. <https://doi.org/10.1186/s12935-017-0455-1> PMID: 29021715
38. Dastsooz H, Cereda M, Donna D, Oliviero S. A Comprehensive Bioinformatics Analysis of UBE2C in Cancers. *Int J Mol Sci.* 2019; 20(9). <https://doi.org/10.3390/ijms20092228> PMID: 31067633
39. Hu J, Wu X, Yang C, Rashid K, Ma C, Hu M, et al. Anticancer effect of icaritin on prostate cancer via regulating miR-381-3p and its target gene UBE2C. *Cancer Med.* 2019; 8(18):7833–45. <https://doi.org/10.1002/cam4.2630> PMID: 31646760
40. Dastsooz H, Cereda M, Donna D, Oliviero S. A Comprehensive Bioinformatics Analysis of UBE2C in Cancers. *International Journal of Molecular Sciences.* 2019; 20(9). <https://doi.org/10.3390/ijms20092228> PMID: 31067633
41. Chen Z, Zhang C, Wu D, Chen H, Rorick A, Zhang X, et al. Phospho-MED1-enhanced UBE2C locus looping drives castration-resistant prostate cancer growth. *EMBO J.* 2011; 30(12):2405–19. <https://doi.org/10.1038/emboj.2011.154> PMID: 21556051
42. Hu R, Lu C, Mostaghel EA, Yegnasubramanian S, Gurel M, Tannahill C, et al. Distinct transcriptional programs mediated by the ligand-dependent full-length androgen receptor and its splice variants in castration-resistant prostate cancer. *Cancer Res.* 2012; 72(14):3457–62. <https://doi.org/10.1158/0008-5472.CAN-11-3892> PMID: 22710436
43. Warren AY, Massie CE, Watt K, Luko K, Orafidiya F, Selth LA, et al. A reciprocal feedback between the PDZ binding kinase and androgen receptor drives prostate cancer. *Oncogene.* 2019; 38(7):1136–50. <https://doi.org/10.1038/s41388-018-0501-z> PMID: 30237440