Pre-Clinical Research Report

Journal of INTERNATIONAL MEDICAL RESEARCH

Journal of International Medical Research
49(6) 1–18

DOI: 10.1177/03000605211016624
journals.sagepub.com/home/imr
SAGE

Identification of crucial genes and pathways associated with prostate cancer in multiple databases

Hanxu Guo[1], Zhichao Zhang[1], Yuhang Wang[1] and Sheng Xue[2]

Abstract

Objective: Prostate cancer (PCa) is a malignant neoplasm of the urinary system. This study aimed to use bioinformatics to screen for core genes and biological pathways related to PCa.

Methods: The GSE5957 gene expression profiles were obtained from the Gene Expression Omnibus (GEO) database to identify differentially expressed genes (DEGs). Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of the DEGs were constructed by R language. Furthermore, protein–protein interaction (PPI) networks were generated to predict core genes. The expression levels of core genes were examined in the Tumor Immune Estimation Resource (TIMER) and Oncomine databases. The cBioPortal tool was used to study the co-expression and prognostic factors of the core genes. Finally, the core genes of signaling pathways were determined using gene set enrichment analysis (GSEA).

Results: Overall, 874 DEGs were identified. Hierarchical clustering analysis revealed that these 24 core genes have significant association with carcinogenesis and development. *LONRF1, CDK1, RPS18, GNB2L1* (*RACK1*), *RPL30*, and *SEC61A1* directly related to the recurrence and prognosis of PCa.

Conclusions: This study identified the core genes and pathways in PCa and provides candidate targets for diagnosis, prognosis, and treatment.

Keywords

Bioinformatics analysis, prostate cancer, differentially expressed genes, biological pathway, The Cancer Genome Atlas, prognostic biomarker, survival

Date received: 24 December 2020; accepted: 14 April 2021

[1]School of Clinical Medicine, Bengbu Medical College, Bengbu, China
[2]Department of Urology, The First Affiliated Hospital of Bengbu Medical College, Bengbu, China

Corresponding author:
Sheng Xue, Department of Urology, the First Affiliated Hospital of Bengbu Medical College, Bengbu 233000, China.
Email: drxuesheng@163.com

Introduction

Prostate cancer (PCa) is currently a major cause of cancer morbidity and mortality.[1] In China, the rates of incidence and mortality of PCa have increased in all age groups from 1990 to 2017, and timely intervention should be done for individuals less than 40 years old.[2] Although prostate-specific antigen (PSA) levels have been widely used to detect PCa, overdiagnosis and overtreatment of PCa are still limitations of such screening methods.[3] Now, numerous dependable and efficient biomarkers have been identified and proved to be prognostic. Therefore, exploring core genes in PCa has important clinical significance.

Because of high-throughput sequencing technology and bioinformatics methods, it is easier to screen for differentially expressed genes (DEGs) to discover inner signaling networks and relationships between genes.[4] Fibroblast growth factor 2 (*FGF2*),[5] cyclin kinase subunit 2 (*CKS2*),[6] RING finger protein 7 (*RNF7*),[7] cyclin-dependent kinase 1 (*CDK1*),[8] and other core genes were screened out based on their significant effects on the development and progression of PCa. However, several studies have indicated that there are clear differences among them. Consequently, potential genes still need to be explored and verified.

Until now, the gene profile GSE5957 had not been reported or previously screened. This dataset contains 25 experiments that were performed for PCa tissues and three for normal prostate (NP) tissues. All samples and common cases were labeled with Cy5-dUTP. In the present study, a total of 874 DEGs were detected between PCa and NP tissues. To predict the core genes and molecular mechanisms, protein–protein interaction (PPI) networks and Cytoscape software were applied, respectively. The Tumor Immune Estimation Resource (TIMER) database, multiple databases from cBioPortal, and the Oncomine public database were used to examine the expression and prognostic value between cancer and normal tissues.

Materials and methods

Microarray data

The GSE5957 dataset, which contains 25 PCa tissue samples and three matched NP tissue samples, were obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). GPL4381 (UGI Human 14112 V1.0) was converted to official gene symbols based on Database for Annotation Visualization and Integrated Discovery (DAVID; https://david.ncifcrf.gov/), which contains 14,112 probes. The construction of the microarrays used in this study was carried out following Brown's method (available at http://cmgm.stanford.edu/pbrown/protocols/index.html). A "genome-wide" cDNA microarray consisting of 14,061 sequences was generated. These included full-length and partial cDNAs representing novel, known, and control genes provided by United Gene Holdings Group Co., Ltd. The genes represented in the arrays were identified based on high similarity to the sequences in the Unigene database of NCBI by performing Blast (http://www.ncbi.nlm.nih.gov/).

Data processing and screening of DEGs

The affy package of the R programing language (Ver. 4.0.2; R Core Team, 2014) was used to read the GSE5957 dataset. Robust Multi-array Average method was preconditioned. The DEGs were screened based on P-value < 0.05 and |log2FC| (fold change) ≥ 1.[9]

Gene ontology (GO) and pathway enrichment analysis

Gene ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment information of DEGs were adopted in DAVID previously. The standards of screening were $P < 0.05$ with count number ≥ 1.5. The visualization of results was generated by R programing language.

PPI network construction and module analyses

The Search Tool for the Retrieval of Interacting Genes (STRING) database online tool (http://string-db.org/)[10] was used to construct the PPI network mapping, which served as a unique resource for further experimentation and analysis leading to the identification of disease-modifier genes and new drug targets.[11] Criteria of degree cutoff ≥ 2, node score cutoff ≥ 2, K-core ≥ 2, and max depth = 100, as well as the MCODE and Centiscape 2.2 App of Cytoscape (Ver. 3.7.2) based on all DEGs,[12,13] were used for network visualization and identification of core genes.

Core gene selection and prognosis analysis

Based on the degrees ≥ 12, the most significant model was established as the PPI network. cBioPortal (http://www.cbioportal.org) (Table 1) was used to present the co-expression network. Afterwards, to qualify the individual prognostic value for PCa, log2 mRNA expression data (The Cancer Genome Atlas Prostate Adenocarcinoma (TCGA PRAD)) were submitted to TIMER (https://cistrome.shinyapps.io/timer/).[14] We analyzed the Kaplan–Meier (KM) curves of the candidate core genes, which are presented with cBioPortal. A P-value < 0.05 was the criterion for statistical significance. Five databases[15–19] are included in cBioPortal, which verified the effects on patient overall survival (OS) in large samples.

External dataset evaluation and verification

Based on Oncomine (http://www.oncomine.com),[20] the interactions between the core genes and metastasis state were validated. To map all human proteins in cells, tissues, and organs and the pathology of core genes on transcriptional and translational levels, the Human Protein Atlas (HPA) (http://www.proteinatlas.org/) was applied to evaluate the core genes. With the clinical data from Oncomine and TCGA, we investigated the mutual relationships between core gene expression levels and clinical stage. Based on the UALCAN (http://ualcan.path.uab.edu/index.html) online tool,[21] we examined the relevance between the core genes and Gleason scores.

Gene set enrichment analysis (GSEA)

GSEA (http://software.broadinstitute.org/gsea/index.jsp)[22] is a productive tool to predict the functional effect of the core genes. The most significant pathways were screened by P-values < 0.05. "ggplot2" packages were applied to visualize the result in R programing language (Ver. 4.0.2).

Results

Screening of DEGs and data processing

Information regarding mRNA expression is included in dataset GSE5957. A total of 6901 official gene symbols were identified and the expression of each gene was established. Limma R package was applied to

Table 1. Detailed results of Module 1 and Module 2.

SUID	MCODE cluster	MCODE node status	MCODE score	Gene name
91	Cluster 2	Clustered	13	*RPL5*
86	Cluster 2	Clustered	13	*RPL18*
84	Cluster 2	Clustered	13	*RPL9*
82	Cluster 2	Clustered	13	*RPL18A*
81	Cluster 2	Clustered	13	*RPL30*
79	Cluster 2	Clustered	13	*MRPL3*
78	Cluster 2	Clustered	13	*MRPL13*
108	Cluster 2	Clustered	13	*EIF1AX*
76	Cluster 2	Clustered	13	*RPS18*
75	Cluster 2	Clustered	13	*GNB2L1*
167	Cluster 2	Clustered	13	*MRPL1*
164	Cluster 2	Clustered	13	*SEC61A1*
419	Cluster 2	Clustered	12	*MTIF2*
193	Cluster 2	Seed	13	*MRPS14*
97	Cluster 2	Clustered	13	*MRPL15*
2271	Cluster 1	Clustered	21	*RBX1*
1887	Cluster 1	Clustered	21	*ANAPC7*
2494	Cluster 1	Clustered	21	*RNF19A*
2493	Cluster 1	Clustered	21	*RNF130*
2301	Cluster 1	Clustered	21	*ASB7*
1945	Cluster 1	Clustered	21	*ATG7*
2295	Cluster 1	Clustered	21	*RNF7*
2103	Cluster 1	Clustered	21	*UFL1*
2772	Cluster 1	Clustered	13	*LONRF1*
2420	Cluster 1	Clustered	21	*UBR2*
1908	Cluster 1	Clustered	21	*CDC27*
2323	Cluster 1	Clustered	21	*FBXW4*
2609	Cluster 1	Clustered	21	*PJA2*
2704	Cluster 1	Clustered	21	*HERC6*
2383	Cluster 1	Clustered	21	*UBE2J1*
1935	Cluster 1	Clustered	21	*CDC16*
1934	Cluster 1	Clustered	21	*ANAPC13*
2630	Cluster 1	Clustered	21	*RNF41*
2374	Cluster 1	Clustered	21	*RNF4*
2469	Cluster 1	Clustered	21	*TRIM37*
2373	Cluster 1	Clustered	21	*UBE2V2*
2788	Cluster 1	Seed	21	*TRIM4*

filter the DEGs (criteria: $P < 0.05$ and $|\log 2$ fold change$| \geq 1$). A total of 874 DEGs were discerned between PCa and NP samples, including 226 upregulated genes and 648 downregulated genes. A cluster heatmap of the 874 DEGs and a volcano plot are presented in Figure 1.

KEGG and GO enrichment of DEGs

The functional classifications of the 874 DEGs were generated using DAVID. Regulation of cell cycle process, regulation of cellular protein localization, and cellular component disassembly of biological

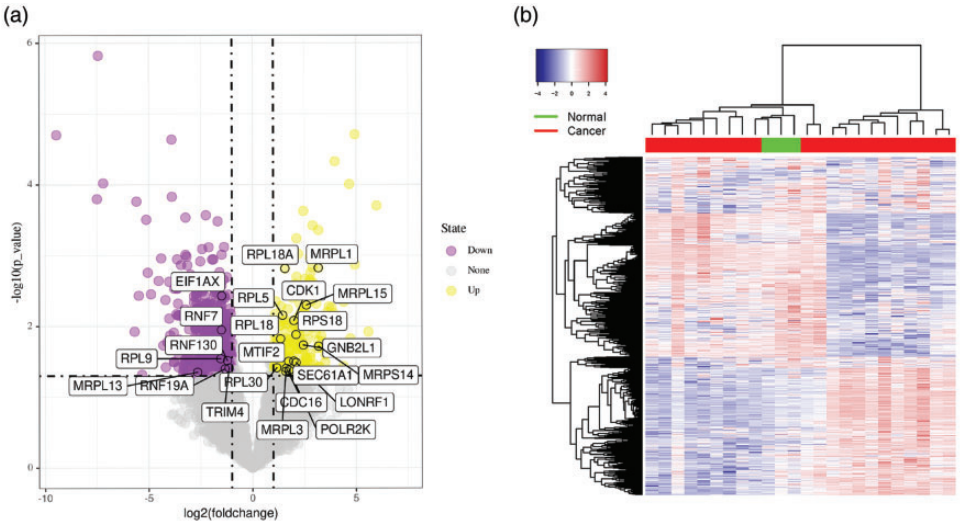

Figure 1. Volcano plot and cluster heatmap of the differentially expressed genes (DEGs). (a) Volcano plot: Significantly upregulated DEGs, significantly downregulated DEGs, and insignificantly changed genes are represented as blue, purple, and gray dots, respectively, in prostate cancer (PCa) and normal prostate (NP) tissues. The top 24 candidate DEGs are noted. (b) Cluster heatmap of the DEGs.

procession were the most significantly enriched groups by GO analysis. For molecular function, DEGs were significantly enriched for actin, ubiquitin-like protein ligase, 3'-5' exonuclease activity, and transcription factor binding. For the cellular component group of genes, mitochondrial envelope, focal adhesion, and dendrite were extremely enriched. KEGG pathway analysis indicated that cell cycle, N-Glycan biosynthesis, and Hippo signaling pathway were largely enriched. (Figure 2).

PPI and module analysis

Cytoscape software and the online database STRING (available online: https://string-db.org/) were used to identify core genes, which consisted of 862 nodes interacting with each other via 3124 edges. Based on the APP of Centiscape and MCODE, the most statistically significant module 14 core genes that were screened are shown in Figure 3 with degreeMCODE scores $\leq$13 (including LONFR1 and all genes in Module 2 (C)) in MCODE (Table 1) and the top 10-degree genes in Centiscape (Table 2). The GO and KEGG enrichment analyses of this module were generated using DAVID. The most significant module 14 core genes were mainly enriched for ribosomal subunit, translational elongation, ubiquitin-dependent protein catabolic process, ribosome, and cell cycle. (Table 3).

Candidate gene selection and survival analysis

Candidate core genes (24 core genes) were obtained from the results of MCODE and Centiscape above. The co-expression network was generated using cBioPortal (Table 4). Data from TCGA were analyzed with the platform TIMER, which was applied to indicate the expression differences of core genes between PCa and NP tissues. As suggested in Figure 4, the mRNA levels of LON Peptidase N-Terminal Domain And Ring Finger 1 (*LONRF1*),

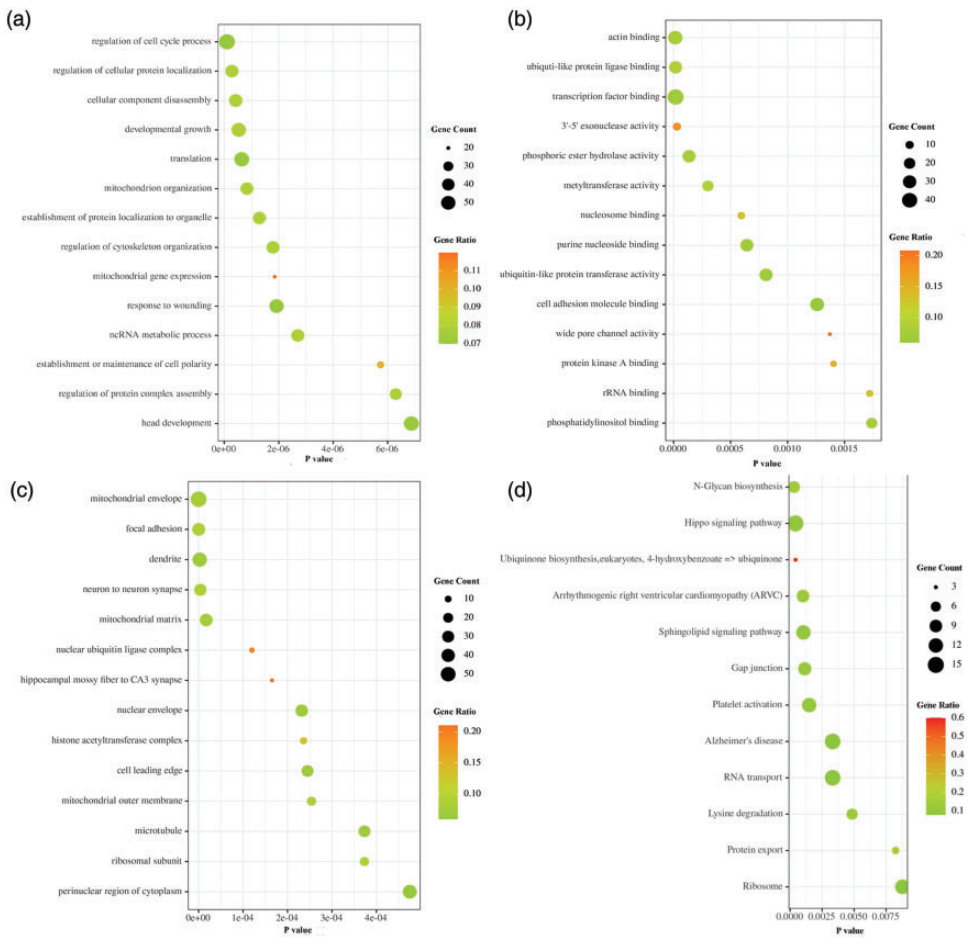

Figure 2. The Gene Ontology (GO)/Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of the differentially expressed genes (DEGs). (a) GO terms of biological process in DEGs. (b) GO terms of molecular function in DEGs. (c) GO terms of cell component in DEGs. (d) KEGG pathway terms in DEGs.

CDK1, Ring Finger Protein 130 (*RNF130*), Ribosomal Protein L18A (*RPL18A*), Mitochondrial ribosomal protein L3 (*MRPL3*), *MRPL13, RPL5, RPL18, RPL9, RPL18A MRPL15,* guanine nucleotide-binding protein subunit beta-2-like 1 (*GNB2L1*/*RACK1*), *RPL18*, *RPL30*, and Protein transport protein Sec61 subunit alpha isoform 1 (*SEC61A1*) had statistically significant values ($P < 0.001$). The prognostic values of the core genes based on multiple databases were evaluated with the cBioPortal online tool. As shown in Figure 5, six of the abovementioned 15 candidate core genes with the lowest *P*-values for both expression and OS results of the above 15 candidate genes demonstrated great prognostic value for PCa patients. Furthermore, we used the Oncomine database to create an overview of core genes in PCa tissues compared with NP tissues (Figure 6).

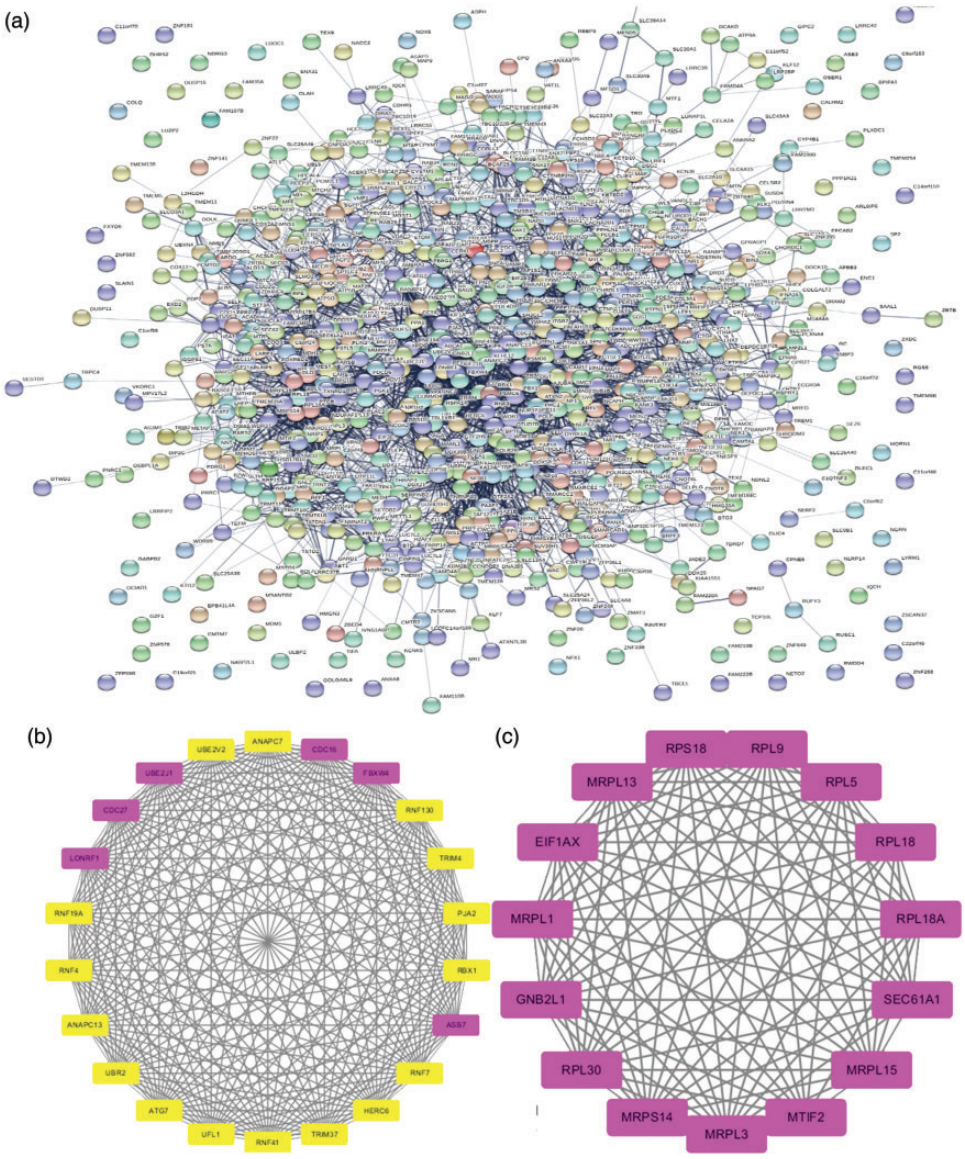

Figure 3. (a) Protein-protein interaction (PPI) network for differentially expressed genes (DEGs). (b) Module 1 of DEGs. (C) Module 2 of DEGs.

Collection of true core genes

Six genes were predicted to be the core factors affecting PCa based on the above analyses. To clarify our predictions on transcriptional and translational levels, HPA databases provided the cases of immunohistochemistry (IHC)[23], which were applied to verify the differential protein expression of key factors. As suggested in Figure 7, the PCa group showed stronger staining than the NP group for upregulated genes. The reverse was true for the downregulated genes. Furthermore, the

Oncomine database was used to identify the mRNA expression differences of core genes between local lesions and metastatic patients. The results indicated that the core genes played a significant role in the carcinogenesis of PCa.

GSEA, clinical correlation

Using the UALCAN online tool, *LONRF1, CDK1, RPL18, GNB2L1 (RACK1), RPL30*, and *SEC61A1* were determined to have close relationships with patients' Gleason scores (Figure 8). Generally, as the Gleason score elevated, the expression levels of the core factors also increased. We then observed that ribosome, oxidative phosphorylation, Parkinson's disease, spliceosome, and Alzheimer's disease were highly enriched with high expression levels of the six core genes, except for *CDK1* (Figure 9). The *P*-values of the different core genes are shown in Table 5.

Table 2. Detailed results of the Centiscape app.

Betweenness	Closeness	Degree	Gene name
36647.74	4.79E-04	57	*CDK1*
24255.41	4.60E-04	54	*RBX1*
18095.66	4.40E-04	42	*POLR2K*
13717.48	4.49E-04	41	*CDC27*
19510.11	4.61E-04	39	*GNB2L1*
8274.128	4.41E-04	38	*RPL5*
17344.4	4.35E-04	38	*BPTF*
9148.067	4.36E-04	37	*RPL9*
7421.087	4.29E-04	36	*CDC16*
18538.91	4.55E-04	36	*PSMC6*

Discussion

Although several studies on the molecular processes and progression of PCa have been performed, PCa remains a common malignant tumor of the genitourinary system in elderly men.[24] A full understanding of the pathogenesis of PCa remains indeterminate. Thus, it is crucial to explore the underlying biomarkers for clinical treatment of PCa to improve patient outcomes.[25]

In this study, 874 (226 upregulated 648 downregulated) DEGs were sorted from the GEO dataset GSE5957. GO/KEGG pathway analyses suggested that the DEGs were enriched for cell cycle, cellular component, and protein localization. The cell cycle

Table 3. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of differentially expressed genes (DEGs) in the most significant module 14 core genes.

Pathway ID	Term	Count	*P*-value
GO:0019787	ubiquitin-like protein transferase activity	16	7.852e-19
GO:0000209	ribosomal subunit	12	8.933e-17
GO:0006414	protein polyubiquitination	13	1.618e-15
GO:0042176	translational elongation	7	1.914e-9
GO:0045116	regulation of protein catabolic process	8	1.599e-7
GO:1903008	protein neddylation	3	2.094e-6
GO:0051865	organelle disassembly	4	2.208e-5
GO:0015935	protein autoubiquitination	3	2.099e-4
GO:0006417	small ribosomal subunit	3	2.183e-4
GO:0031330	regulation of translation	5	6.792e-4
hsa03010	Ribosome	11	5.984e-16
hsa04120	Ubiquitin mediated proteolysis	8	4.315e-11
hsa04141	Protein processing in endoplasmic reticulum	3	2.000e-3

The genes were mainly enriched for celubiquitin-like protein transferase activity, ribosomal subunit, protein polyubiquitination, ubiquitin-mediated proteolysis, and protein processing in the endoplasmic reticulum.

Table 4. Interaction network between the hub genes and their co-expression genes using cBioPortal.

Gene	Co-expression gene	Log_2OR	*P*-Value	Gene	Co-expression Gene	Log_2OR	*P*-Value
CDK1	*RPS18*	2.824	<0.001	*RPL30*	*MRPS14*	1.963	<0.001
CDK1	*RPL9*	2.753	<0.001	*RPL30*	*MRPL1*	1.616	0.004
CDK1	*MTIF2*	2.553	<0.001	*RPL30*	*MTIF2*	1.347	0.027
CDK1	*MRPL3*	2.464	<0.001	*RPS18*	*MRPS14*	>3	<0.001
CDK1	*SEC61A1*	2.392	<0.001	*RPS18*	*MRPL15*	1.825	0.005
CDK1	*RPL18*	2.335	0.006	*RPS18*	*SEC61A1*	1.724	0.016
CDK1	*RACK1*	2.225	<0.001	*RPS18*	*MTIF2*	2.32	0.029
CDK1	*RPL18A*	2.211	0.018	*LONRF1*	*MTIF2*	2.284	<0.001
CDK1	*MRPS14*	2.079	0.002	*LONRF1*	*MRPS14*	1.961	<0.001
CDK1	*CDC27*	1.743	0.006	*LONRF1*	*RPL18*	2.463	<0.001
CDK1	*POLR2K*	1.093	0.004	*RPS18*	*MRPL15*	1.825	0.005
RACK1	*RPS18*	>3	<0.001	*LONRF1*	*TRIM4*	1.398	0.01
RACK1	*MRPL3*	2.325	<0.001	*LONRF1*	*SEC61A1*	1.188	0.011
RACK1	*RPL18*	2.933	0.001	*LONRF1*	*EIF1AX*	1.349	0.012
RACK1	*MRPS14*	2.171	0.004	*LONRF1*	*CDK1*	1.181	0.014
RACK1	*RPL18A*	2.801	0.005	*RPS18*	*SEC61A1*	1.724	0.016
RPL30	*MRPL13*	>3	<0.001	*LONRF1*	*RPL9*	1.884	0.02
RPL30	*MRPL15*	>3	<0.001	*RPS18*	*MTIF2*	2.32	0.029
RPL30	*SEC61A1*	2.218	<0.001	*SEC61A1*	*MRPL15*	2.091	<0.001
RPL30	*EIF1AX*	2.037	<0.001	*SEC61A1*	*MTIF2*	2.928	<0.001
RPL30	*MRPL3*	1.702	<0.001	*SEC61A1*	*MRPS14*	1.948	0.003

progression score (CCP) has been well demonstrated to be better than existing assessments for elucidating the potential aggressiveness of PCa in an individual.[26] Multiple studies have reported that some genes can inhibit cell cycle progression from G0/G1 to S phase, as well as be regulated by certain microRNAs in PCa.[27–29] In androgen receptor (AR) signaling, which plays a crucial role in the development and progression of PCa, some RNAs can activate or inactivate AR signaling to inhibit the cell cycle of PCa cells.[30,31] Different proteins with various cell localizations can affect PCa tumorigenesis, such as Reticulon 4 (RTN4),[32] which is a reticulon family protein localized in the endoplasmic reticulum. In addition, the overexpression of N-Glycan in PCa is a possible biomarker for diagnosis.[33] Additionally, the hippo pathway, which was found to be enriched by KEGG analysis, has a strong association with regulating PCa development.[34] Furthermore, identification of unknown immune editing components helps prostate tumor cells escape from immune surveillance.[35] Overall, these biological functions and pathways have close correlations with the development and progression of PCa.

A PPI network was constructed with DEGs and a total of 24 candidates with degree ≥ 20 are listed in Table 5. We identified *LONRF1, CDK1, RPL18, GNB2L1 (RACK1), RPL30*, and *SEC61A1* as real core genes by five databases (included TCGA). These genes have significant prognostic value. It is reported that in the absence of interphase CDKs, CDK1 executes all the events that are required to drive cell division,[36] which may lead to the proliferation of cancer cells. Xie et al.[37] indicated that CDK1 controlled mitochondrial metabolism, which plays key roles in the flexible bioenergetics required for tumor cell survival. Furthermore, the

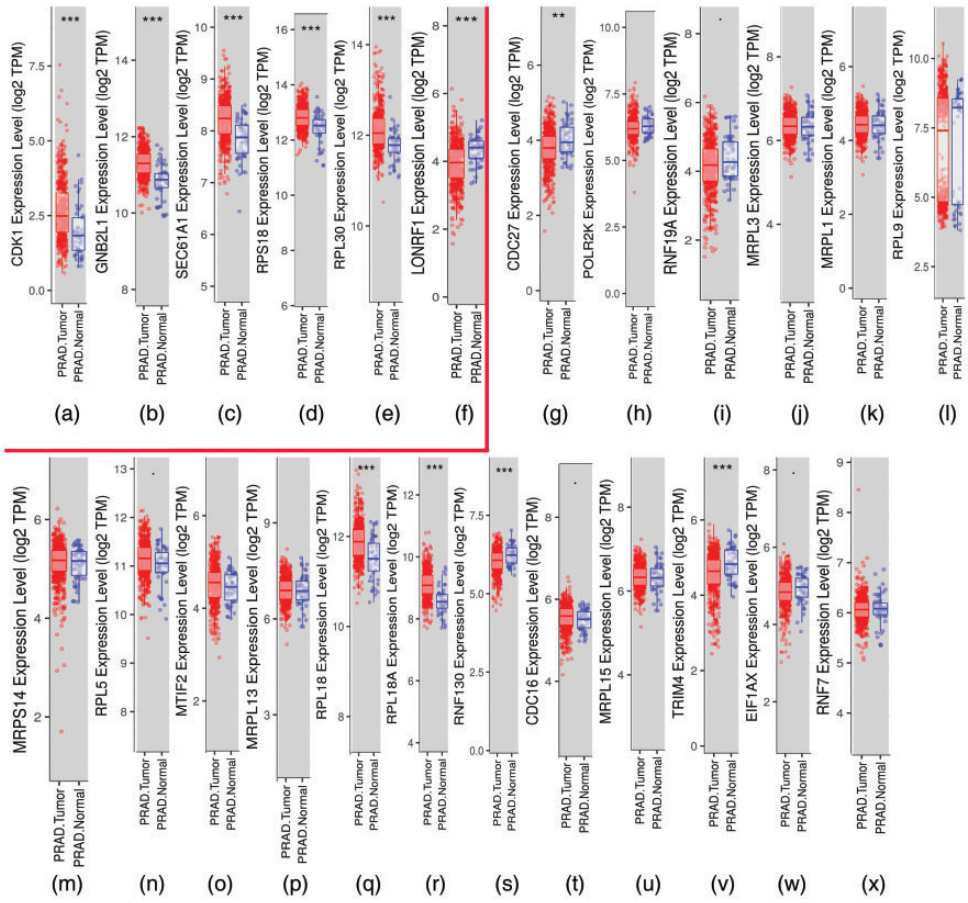

Figure 4. Differential analysis of expression between prostate cancer (PCa) and normal prostate (NP) tissues by Tumor Immune Estimation Resource (TIMER). (a–f) Box plots of core genes (LON Peptidase N-Terminal Domain And Ring Finger 1 (LONRF1), Cyclin Dependent Kinase 1 (CDK1), Ribosomal Protein S18 (RPS18), guanine nucleotide-binding protein subunit beta-2-like 1 (GNB2L1), Ribosomal Protein L30 (RPL30), and Protein transport protein Sec61 subunit alpha isoform 1 (SEC61A1)) presented by TIMER. The six core genes were found to be risk factors in The Cancer Genome Atlas (TCGA) database. (G–X) Box plots of the other 18 candidate genes.

overexpression of CDK1 correlates with poor prognosis and metastasis in various adenocarcinomas, such as pancreatic cancer and breast cancer.[38,39] Several inhibitors targeting CDK1 have been tested in clinical trials (NCT00141297, NCT00147485, and NCT00292864) for certain human neoplasias.[40] CDK1 is essential for cell migration and accumulating evidence now demonstrates multiple roles for RACK1 (GNB2L1) in regulating migration and invasion of tumor cells.[41] RACK1 has been recognized as an independent biomarker for poor clinical outcome in pancreatic and breast cancers.[42,43] In PCa cells, high expression of RACK1 promotes cell growth and metastasis *in vivo*.[41,44] Transient Receptor Potential Cation Channel Subfamily M Member 8 (TRPM8) promotes hypoxic growth of

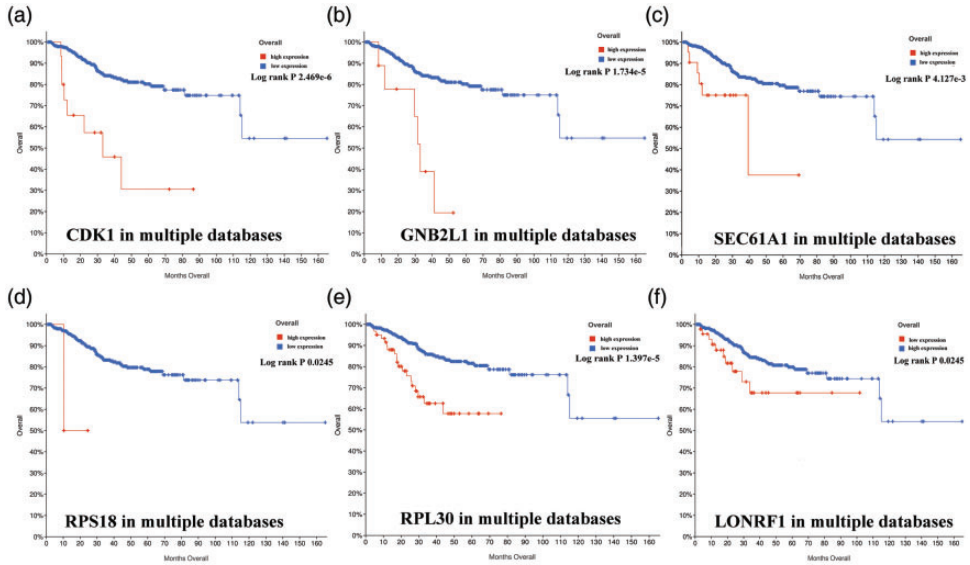

Figure 5. Overall survival analyses of prostate cancer (PCa) patients based on high or low expression of six core genes from multiple databases. Survival curves for (a) Cyclin Dependent Kinase 1 (CDK1), (b) guanine nucleotide-binding protein subunit beta-2-like 1 (GNB2L1), (c) Protein transport protein Sec61 subunit alpha isoform 1 (SEC61A1), (D) Ribosomal Protein S18 (RPS18), (e) Ribosomal Protein L30 (RPL30), and (f) LON Peptidase N-Terminal Domain And Ring Finger 1 (LONRF1), are shown. Log-rank *P*-values < 0.05 were considered to be statistically significant.

PCa cells via RACK1-mediated stabilization of Hypoxia-inducible factor (HIF)-1α.[45] High expression of RACK1 also leads to PCa cell growth. RPS18 is a type of ribosomal protein that recently has been determined to promote breast cancer metastasis.[46] This protein was first discovered with the structure of RNA in 1991, but there are few studies regarding its molecular functions in cells. Upregulation of RPS18 in PCa has great statistical importance, as indicated by TIMER. However, in the Oncomine database, high expression (Grasso Prostate) and low expression (Tomlins Prostate) of RPS18 both have significant *P*-values in PCa. Furthermore, we suggest that RPS18 overexpression increases ribosomal content and global translation in PCa cells compared with RPL15. RPL30 maintains cell growth and survival, and it has been associated with outcomes in medulloblastoma.[47] Moreover, methylation of RPL30 downregulated the dysfunction of anti-apoptosis that is involved in non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH) pathways, which could lead to hepatocarcinogenesis.[48] In head and neck squamous cell carcinoma, *RPL30* and other genes[49] were determined as the internal core genes with stable expression. *SEC61A1*, as a hallmark DNA repair gene, was identified as a potential independent indicator of prognosis in hepatocellular carcinoma in a gene-wide association study.[50] Beyond carcinoma, mutations and loss-of-function of *SEC61A1* can cause multiple myeloma and kidney disease.[51,52] For *LONRF1*, the only downregulated gene, there are no existing reports on any correlations with cancer. However, according to the

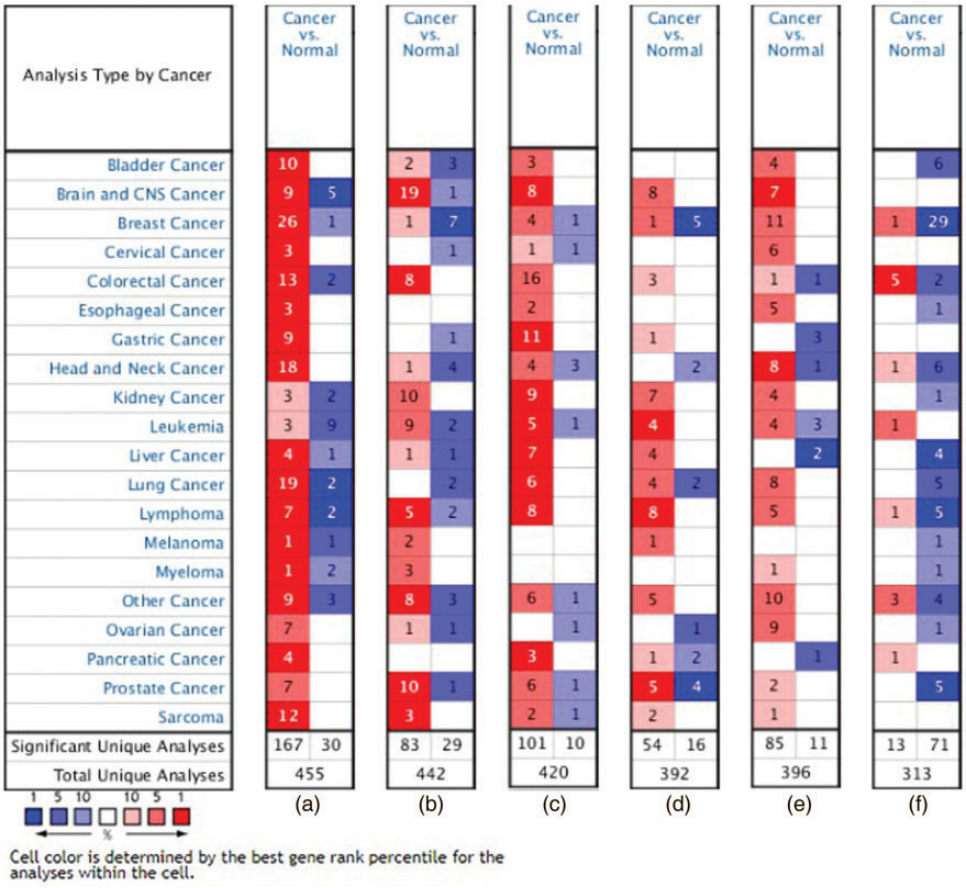

Figure 6. Based on Oncomine data, overviews of mRNA levels of core genes in various types of cancer were constructed. Genes include (a) Cyclin Dependent Kinase 1 (CDK1), (b) guanine nucleotide-binding protein subunit beta-2-like 1 (GNB2L1), (c) Ribosomal Protein L30 (RPL30), (d) Ribosomal Protein S18 (RPS18), (e) Protein transport protein Sec61 subunit alpha isoform 1 (SEC61A1), and (f) LON Peptidase N-Terminal Domain And Ring Finger 1 (LONRF1). Statistically significant mRNA overexpression or low expression of genes are presented in red and blue, respectively. Grid color was determined by the best gene rank percentile for analysis within the grids. The threshold settings: gene rank percentile = top 10%, $P = 0.05$, and fold change (FC) = All.

Oncomine data and OS analysis, the significantly high expression of this gene in normal tissue samples has great prognostic value. Consequently, there is much potential for research on whether four genes (*RPS18, RPL30, SEC61A1*, and *LONRF1*) play a key role in the pathogenesis of PCa. As for CDK1, one study based on the GEO single database indicated that CDK1 levels were significantly higher in PCa tissues compared with normal tissues.[53]

CDC27, POLR2K, RNF19A, and *RNF7* were several of the 20 candidate genes that were demonstrated to be risk factors of PCa. For *CDC27*, reports found that the novel *CDC27-OAT* gene fusion was present in PCa patients.[54] CDC27 has not

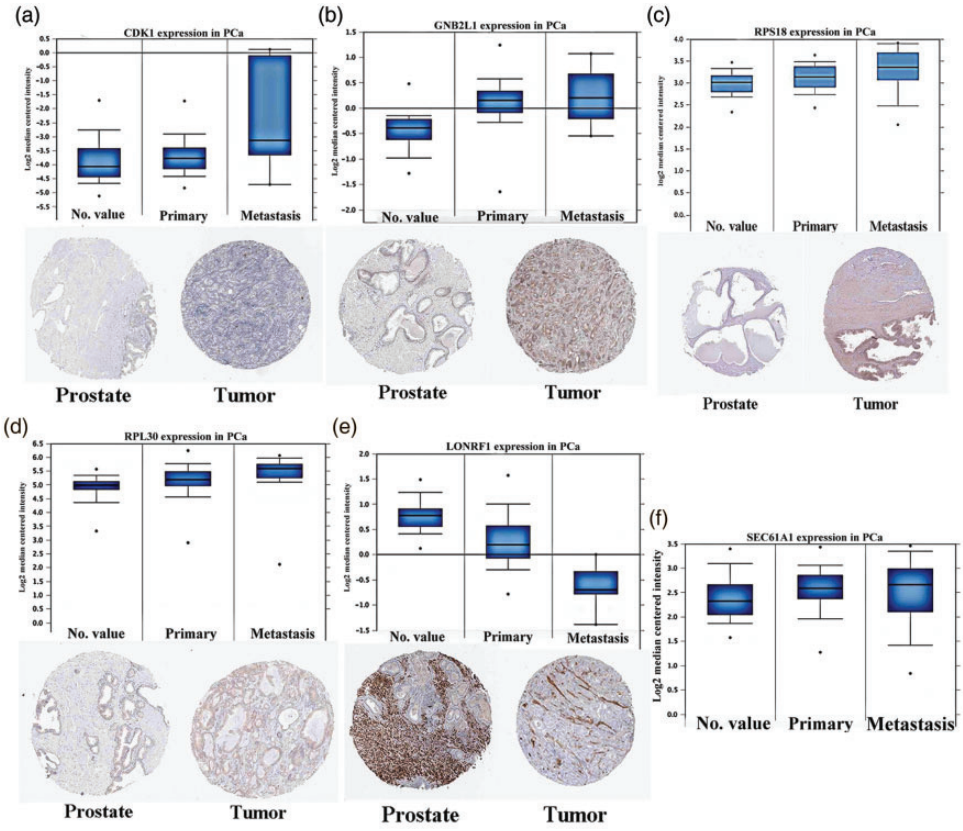

Figure 7. Gene expression of core genes on transcriptional and translational levels. Verification of the expression of core genes on transcriptional and translational level was done with the Oncomine database and the Human Protein Atlas database (immunohistochemistry (IHC)), respectively. Transcriptional data are shown for normal prostate (NP) tissues, primary prostate cancer (PCa) tumors, and metastatic PCa tumors. IHC data are shown for NP tissues and PCa tumors. Genes include (a) Cyclin Dependent Kinase 1 (CDK1), (b) guanine nucleotide-binding protein subunit beta-2-like 1 (GNB2L1), (c) Ribosomal Protein S18 (RPS18), (d) Ribosomal Protein L30 (RPL30), (e) LON Peptidase N-Terminal Domain And Ring Finger 1 (LONRF1), and (f) Protein transport protein Sec61 subunit alpha isoform 1 (SEC61A1).

previously been reported to be significantly mutated in PCa. Such a mutation landscape suggests that investigating the biological functions of these cancer cells is quite important.[55] Additionally, *POLR2K* and *APT6V1A* were screened out between PCa tissues and normal tissues as novel DEGs. They both play important roles in supporting other pathways that facilitate tumor growth.[56] In our study, we also found that *POLR2K* and *EIF1AX* had significant prognostic value for OS (Supplementary Figure) of PCa. RNF19A was identified as an RNA transcript that is present at significantly higher levels in the blood of PCa patients compared with healthy controls and is a potential clinically relevant biomarker for PCa early detection.[57] RNF7 was originally identified as a redox-inducible antioxidant protein.[58] One study demonstrated that RNF7 knockdown inhibited PCa cell proliferation and tumorigenesis, suggesting that RNF7 may be a promising target for PCa treatment.[7]

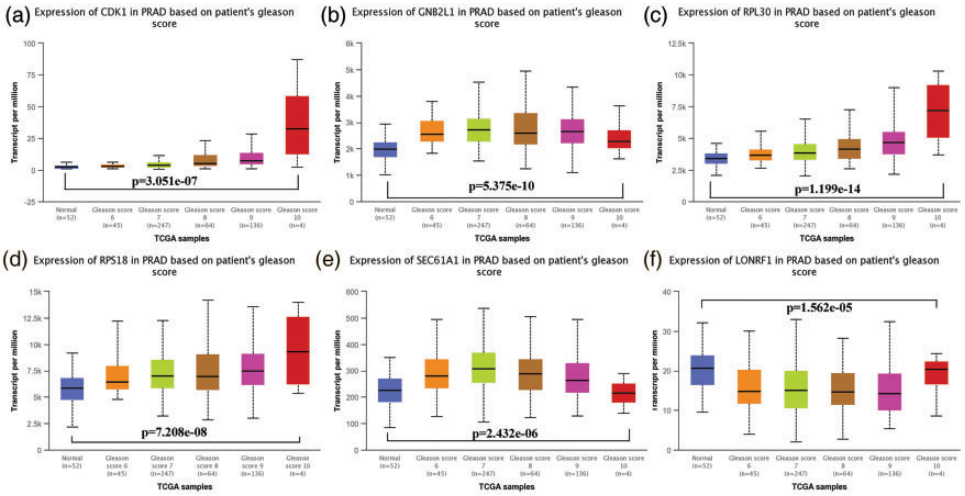

Figure 8. Gene expression of core genes represented by different Gleason scores. The gene expression of (a) Cyclin Dependent Kinase 1 (CDK1), (b) guanine nucleotide-binding protein subunit beta-2-like 1 (GNB2L1), (c) Ribosomal Protein L30 (RPL30), (d) Ribosomal Protein S18 (RPS18), (e) Protein transport protein Sec61 subunit alpha isoform 1 (SEC61A1), (f) LON Peptidase N-Terminal Domain And Ring Finger 1 (LONRF1) represented by different Gleason scores. The results were evaluated by variance analysis.

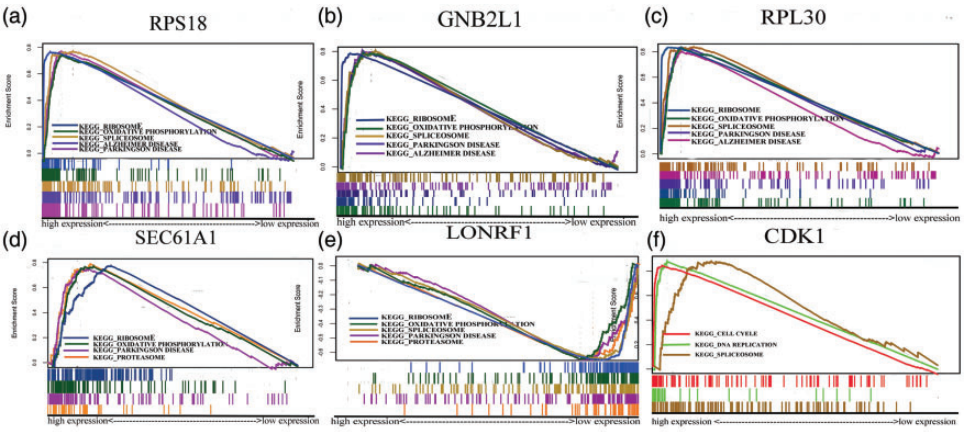

Figure 9. Gene set enrichment analysis (GSEA) of various pathways for core genes. GSEA pathway enrichment analysis of high expression groups of (a) Ribosomal Protein S18 (RPS18), (b) guanine nucleotide-binding protein subunit beta-2-like 1 (GNB2L1), (c) Ribosomal Protein L30 (RPL30), (d) Protein transport protein Sec61 subunit alpha isoform 1 (SEC61A1), (e) LON Peptidase N-Terminal Domain And Ring Finger 1 (LONRF1), and (f) Cyclin Dependent Kinase 1 (CDK1). Pathways with nominal (NOM) $P < 0.05$ and false discovery rate (FDR) $q < 0.06$ were considered to be statistically significant.

In our present study, we identified *LONRF1, CDK1, RPL18, GNB2L1* (*RACK1*), *RPL30*, and *SEC61A1* as crucial components in the diagnosis and treatment of PCa. Research into the associations between molecules in ribosome, oxidative phosphorylation, Parkinson's disease, spliceosome, and Alzheimer's disease could

Table 5. Gene set enrichment analysis (GSEA) results for core genes in The Cancer Genome Atlas Prostate Adenocarcinoma (TCGA PRAD).

Gene	Ribosome		Oxidative phosphorylation		Parkinson's disease	
	NOM *P*-value	FDR q-value	NOM *P*-value	FDR q-value	NOM *P*-value	FDR q-value
GNB2LA	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
RPL30	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
RPS18	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
SEC61A1	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
LONRF1	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
CDK1	/	/	/	/	/	/

The *P*-values of the ribosome, oxidative phosphorylation, and Parkinson's disease for the five core genes are listed.

reveal the underlying causes of cancer and provide novel ideas for identifying target drugs.

Conclusion

Through bioinformatic analyses including GO/KEGG enrichment, PPI networks, core gene identification and validation, module analysis, and GSEA, the present study qualified *LONRF1, CDK1, RPS18, GNB2L1* (*RACK1*), *RPL30*, and *SEC61A1* as sufficient and reliable molecular biomarkers for the diagnosis of PCa. More experimental studies are needed, however, to verify the mechanisms of these genes in PCa.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by Major Online Teaching Reform Program of Bengbu Medical College (No. 2020zdxsjg006) and National University Student Innovation Experimental Project (No. 202010367035).

ORCID iD

Hanxu Guo https://orcid.org/0000-0001-8004-7965

Supplemental material

Supplementary material for this article is available online.

References

1. De Bono JS, Guo C, Gurel B, et al. Prostate carcinogenesis: inflammatory storms. *Nat Rev Cancer* 2020; 20: 455–469. DOI: 10.1038/s41568-020-0267-9.
2. Liu X, Yu C, Bi Y, et al. Trends and age-period-cohort effect on incidence and mortality of prostate cancer from 1990 to 2017 in China. *Public Health* 2019; 172: 70–80. DOI: 10.1016/j.puhe.2019.04.016.
3. Inahara M, Suzuki H, Kojima S, et al. Improved prostate cancer detection using systematic 14-core biopsy for large prostate glands with normal digital rectal examination findings. *Urology* 2006; 68: 815–819. DOI: 10.1016/j.urology.2006.05.010.
4. Clough E and Barrett T. The Gene Expression Omnibus Database. *Methods Mol Biol* 2016; 1418: 93–110. DOI: 10.1007/978-1-4939-3578-9_5.
5. Liu S, Wang W, Zhao Y, et al. 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: 809. DOI: 10.3389/fonc.2020.00809.
6. Wang Y, Wang J, Yan K, et al. Identification of core genes associated with prostate cancer progression and outcome via bioinformatics analysis in multiple

databases. *PeerJ* 2020; 8: e8786. DOI: 10.7717/peerj.8786.
7. Xiao Y, Jiang Y, Song H, et al. RNF7 knockdown inhibits prostate cancer tumorigenesis by inactivation of ERK1/2 pathway. *Sci Rep* 2017; 7: 43683. DOI: 10.1038/srep43683.
8. Luo C, Chen J and Chen L. Exploration of gene expression profiles and immune microenvironment between high and low tumor mutation burden groups in prostate cancer. *Int Immunopharmacol* 2020; 86: 106709. DOI: 10.1016/j.intimp.2020.106709.
9. Ritchie M, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43: e47. DOI: 10.1093/nar/gkv007.
10. Von Mering C, Huynen M, Jaeggi D, et al. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res* 2003; 31: 258–261. DOI: 10.1093/nar/gkg034.
11. Stelzl U, Worm U, Lalowski M, et al. A human protein-protein interaction network: a resource for annotating the proteome. *Cell* 2005; 122: 957–968. DOI: 10.1016/j.cell.2005.08.029.
12. Bandettini WP, Kellman P, Mancini C, et al. MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. *J Cardiovasc Magn Reson* 2012; 14: 83. DOI: 10.1186/1532-429x-14-83.
13. Scardoni G, Petterlini M and Laudanna C. Analyzing biological network parameters with CentiScaPe. *Bioinformatics* 2009; 25: 2857–2859. DOI: 10.1093/bioinformatics/btp517.
14. Li T, Fan J, Wang B, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res* 2017; 77: e108–e110. DOI: 10.1158/0008-5472.Can-17-0307.
15. Abida W, Cyrta J, Heller G, et al. Genomic correlates of clinical outcome in advanced prostate cancer. *Proc Natl Acad Sci U S A* 2019; 116: 11428–11436. DOI: 10.1073/pnas.1902651116.
16. Baca S, Prandi D, Lawrence M, et al. Punctuated evolution of prostate cancer genomes. *Cell* 2013; 153: 666–677. DOI: 10.1016/j.cell.2013.03.021.
17. Kumar A, Coleman I, Morrissey C, et al. Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. *Nat Med* 2016; 22: 369–378. DOI: 10.1038/nm.4053.
18. Armenia J, Wankowicz S, Liu D, et al. The long tail of oncogenic drivers in prostate cancer. *Nat Genet* 2018; 50: 645–651. DOI: 10.1038/s41588-018-0078-z.
19. Hoadley K, Yau C, Hinoue T, et al. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* 2018; 173: 291–304. e296. DOI: 10.1016/j.cell.2018.03.022.
20. Grasso C, Wu Y, Robinson D, et al. The mutational landscape of lethal castration-resistant prostate cancer. *Nature* 2012; 487: 239–243. DOI: 10.1038/nature11125.
21. Chandrashekar D, Bashel B, Balasubramanya S, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* 2017; 19: 649–658. DOI: 10.1016/j.neo.2017.05.002.
22. Subramanian A, Tamayo P, Mootha V, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005; 102: 15545–15550. DOI: 10.1073/pnas.0506580102.
23. Uhlen M, Zhang C, Lee S, et al. A pathology atlas of the human cancer transcriptome. *Science* 2017; 357: eaan2507. DOI: 10.1126/science.aan2507.
24. Siegel R, Miller K and Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019; 69: 7–34. DOI: 10.3322/caac.21551.
25. Intasqui P, Bertolla RP and Sadi MV. Prostate cancer proteomics: clinically useful protein biomarkers and future perspectives. *Expert Rev Proteomics* 2018; 15: 65–79. DOI: 10.1080/14789450.2018.1417846.
26. Sommariva S, Tarricone R, Lazzeri M, et al. Prognostic Value of the Cell Cycle Progression Score in Patients with Prostate Cancer: A Systematic Review and Meta-

analysis. *Eur Urol* 2016; 69: 107–115. DOI: 10.1016/j.eururo.2014.11.038.
27. Long B, Li N, Xu XX, et al. Long noncoding RNA LOXL1-AS1 regulates prostate cancer cell proliferation and cell cycle progression through miR-541-3p and CCND1. *Biochem Biophys Res Commun* 2018; 505: 561–568. DOI: 10.1016/j.bbrc.2018.09.160.
28. Zhang W, Wang S, Zhang X, et al. Transmembrane Channel-Like 5 (TMC5) promotes prostate cancer cell proliferation through cell cycle regulation. *Biochimie* 2019; 165: 115–122. DOI: 10.1016/j.biochi.2019.07.017.
29. Roberto D, Klotz LH and Venkateswaran V. Cannabinoid WIN 55,212-2 induces cell cycle arrest and apoptosis, and inhibits proliferation, migration, invasion, and tumor growth in prostate cancer in a cannabinoid-receptor 2 dependent manner. *Prostate* 2019; 79: 151–159. DOI: 10.1002/pros.23720.
30. Dai X, Liu L, Liang Z, et al. Silencing of lncRNA MALAT1 inhibits cell cycle progression via androgen receptor signaling in prostate cancer cells. *Pathol Res Pract* 2019; 215: 712–721. DOI: 10.1016/j.prp.2019.01.011.
31. Zhang Y, Pitchiaya S, Cieślik M, et al. Analysis of the androgen receptor-regulated lncRNA landscape identifies a role for ARLNC1 in prostate cancer progression. *Nat Genet* 2018; 50: 814–824. DOI: 10.1038/s41588-018-0120-1.
32. Zhao H, Su W, Zhu C, et al. Cell fate regulation by reticulon-4 in human prostate cancers. *J Cell Physiol* 2019; 234: 10372–10385. DOI: 10.1002/jcp.27704.
33. Vermassen T, De Bruyne S, Himpe J, et al. N-Linked Glycosylation and Near-Infrared Spectroscopy in the Diagnosis of Prostate Cancer. *Int J Mol Sci* 2019; 20: 1592. DOI: 10.3390/ijms20071592.
34. Salem O and Hansen CG. The Hippo Pathway in Prostate Cancer. *Cells* 2019; 8: 370. DOI: 10.3390/cells8040370.
35. Su F, Zhang W, Zhang D, et al. Spatial Intratumor Genomic Heterogeneity within Localized Prostate Cancer Revealed by Single-nucleus Sequencing. *Eur Urol* 2018; 74: 551–559. DOI: 10.1016/j.eururo.2018.06.005.
36. Santamaría D, Barrière C, Cerqueira A, et al. Cdk1 is sufficient to drive the mammalian cell cycle. *Nature* 2007; 448: 811–815. DOI: 10.1038/nature06046.
37. Xie B, Wang S, Jiang N, et al. Cyclin B1/CDK1-regulated mitochondrial bioenergetics in cell cycle progression and tumor resistance. *Cancer Lett* 2019; 443: 56–66. DOI: 10.1016/j.canlet.2018.11.019.
38. Piao J, Zhu L, Sun J, et al. High expression of CDK1 and BUB1 predicts poor prognosis of pancreatic ductal adenocarcinoma. *Gene* 2019; 701: 15–22. DOI: 10.1016/j.gene.2019.02.081.
39. Zhang X, Pan Y, Fu H, et al. Nucleolar and Spindle Associated Protein 1 (NUSAP1) Inhibits Cell Proliferation and Enhances Susceptibility to Epirubicin In Invasive Breast Cancer Cells by Regulating Cyclin D Kinase (CDK1) and DLGAP5 Expression. *Med Sci Monit* 2018; 24: 8553–8564. DOI: 10.12659/msm.910364.
40. Malumbres M and Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* 2009; 9: 153–166. DOI: 10.1038/nrc2602.
41. Duff D and Long A. Roles for RACK1 in cancer cell migration and invasion. *Cell Signal* 2017; 35: 250–255. DOI: 10.1016/j.cellsig.2017.03.005.
42. Han H, Wang D, Yang M, et al. High expression of RACK1 is associated with poor prognosis in patients with pancreatic ductal adenocarcinoma. *Oncol Lett* 2018; 15: 2073–2078. DOI: 10.3892/ol.2017.7539.
43. Cao XX, Xu JD, Liu XL, et al. RACK1: A superior independent predictor for poor clinical outcome in breast cancer. *Int J Cancer* 2010; 127: 1172–1179. DOI: 10.1002/ijc.25120.
44. Shen F, Yan C, Liu M, et al. RACK1 promotes prostate cancer cell proliferation, invasion and metastasis. *Mol Med Rep* 2013; 8: 999–1004. DOI: 10.3892/mmr.2013.1612.
45. Yu S, Xu Z, Zou C, et al. Ion channel TRPM8 promotes hypoxic growth of prostate cancer cells via an O2 -independent and RACK1-mediated mechanism of HIF-1α

stabilization. *J Pathol* 2014; 234: 514–525. DOI: 10.1002/path.4413.
46. Ebright RY, Lee S, Wittner BS, et al. Deregulation of ribosomal protein expression and translation promotes breast cancer metastasis. *Science* 2020; 367: 1468–1473. DOI: 10.1126/science.aay0939.
47. Yu L, Guan YJ, Gao Y, et al. Rpl30 and Hmgb1 are required for neurulation in golden hamster. *Int J Neurosci* 2009; 119: 1076–1090. DOI: 10.1080/00207450802330504.
48. Li CW, Chiu YK and Chen BS. Investigating Pathogenic and Hepatocarcinogenic Mechanisms from Normal Liver to HCC by Constructing Genetic and Epigenetic Networks via Big Genetic and Epigenetic Data Mining and Genome-Wide NGS Data Identification. *Dis Markers* 2018; 2018: 8635329. DOI: 10.1155/2018/8635329.
49. Palve V, Pareek M, Krishnan NM, et al. A minimal set of internal control genes for gene expression studies in head and neck squamous cell carcinoma. *PeerJ* 2018; 6: e5207. DOI: 10.7717/peerj.5207.
50. Li N, Zhao L, Guo C, et al. Identification of a novel DNA repair-related prognostic signature predicting survival of patients with hepatocellular carcinoma. *Cancer Manag Res* 2019; 11: 7473–7484. DOI: 10.2147/cmar.S204864.
51. Schubert D, Klein MC, Hassdenteufel S, et al. Plasma cell deficiency in human subjects with heterozygous mutations in Sec61 translocon alpha 1 subunit (SEC61A1). *J Allergy Clin Immunol* 2018; 141: 1427–1438. DOI: 10.1016/j.jaci.2017.06.042.
52. Bolar NA, Golzio C, Živná M, et al. Heterozygous Loss-of-Function SEC61A1 Mutations Cause Autosomal-Dominant Tubulo-Interstitial and Glomerulocystic Kidney Disease with Anemia. *Am J Hum Genet* 2016; 99: 174–187. DOI: 10.1016/j.ajhg.2016.05.028.
53. Deng L, Gu X, Zeng T, et al. Identification and characterization of biomarkers and their functions for docetaxel-resistant prostate cancer cells. *Oncol Lett* 2019; 18: 3236–3248. DOI: 10.3892/ol.2019.10623.
54. Lindquist KJ, Paris PL, Hoffmann TJ, et al. Mutational Landscape of Aggressive Prostate Tumors in African American Men. *Cancer Res* 2016; 76: 1860–1868. DOI: 10.1158/0008-5472.Can-15-1787.
55. Lindberg J, Mills IG, Klevebring D, et al. The mitochondrial and autosomal mutation landscapes of prostate cancer. *Eur Urol* 2013; 63: 702–708. DOI: 10.1016/j.eururo.2012.11.053.
56. Kelly RS, Sinnott JA, Rider JR, et al. The role of tumor metabolism as a driver of prostate cancer progression and lethal disease: results from a nested case-control study. *Cancer Metab* 2016; 4: 22. DOI: 10.1186/s40170-016-0161-9.
57. Bai VU, Hwang O, Divine GW, et al. Averaged differential expression for the discovery of biomarkers in the blood of patients with prostate cancer. *PLoS One* 2012; 7: e34875. DOI: 10.1371/journal.pone.0034875.
58. Duan H, Wang Y, Aviram M, et al. SAG, a novel zinc RING finger protein that protects cells from apoptosis induced by redox agents. *Mol Cell Biol* 1999; 19: 3145–3155. DOI: 10.1128/mcb.19.4.3145.