



OPEN

Network models of prostate cancer immune microenvironments identify ROMO1 as heterogeneity and prognostic marker

Lei Wang^{1,2,6}✉, Xudong Liu^{1,6}, Zhe Liu^{3,6}, Yafan Wang^{1,6}, Mengdi Fan¹, Jinyue Yin¹, Yu Zhang¹, Ying Ma¹, Jia Luo¹, Rui Li¹, Xue Zhao¹, Peiju Zhang¹, Lijun Zhao^{1,2}, Jinke Fan¹, Yuxuan Chen⁴, Wei Lu⁵✉ & Xinqiang Song^{1,2}✉

Prostate cancer (PCa) is the fifth leading cause of death from cancer in men worldwide. Its treatment remains challenging due to the heterogeneity of the tumor, mainly because of the lack of effective and targeted prognostic markers at the system biology level. First, the data were retrieved from TCGA dataset, and valid samples were obtained by consistent clustering and principal component analysis; next, key genes were analyzed for prognosis of PCa using WGCNA, MEGENA, and LASSO Cox regression model analysis, while key genes were screened based on disease-free survival significance. Finally, TIMER data were selected to explore the relationship between genes and tumor immune infiltration, and GSCAlite was used to explore the small-molecule targeted drugs that act with them. Here, we used tumor subtype analysis and an energetic co-expression network algorithm of WGCNA and MEGENA to identify a signal dominated by the *ROMO1* to predict PCa prognosis. Cox regression analysis of *ROMO1* was an independent influence, and the prognostic value of this biomarker was validated in the training set, the validated data itself, and external data, respectively. This biomarker correlates with tumor immune infiltration and has a high degree of infiltration, poor prognosis, and strong correlation with CD8+T cells. Gene function annotation and other analyses also implied a potential molecular mechanism for *ROMO1*. In conclusion, we putative *ROMO1* as a portal key prognostic gene for the diagnosis and prognosis of PCa, which provides new insights into the diagnosis and treatment of PCa.

Abbreviations

PCa	Prostate cancer
WGCNA	Weighted Gene Co-expression Network Analysis
MEGENA	Multiscale Embedded Gene Co-expression Network Analysis
DEG	Differentially expressed genes
PSA	Serum prostate-specific antigen
DFS	Disease-free survival
PFN	Planar filtered network
MCA	Multiscale Clustering Analysis
HPA	The Human Protein Atlas
MHA	Multiscale hub analysis
CTA	Cluster-trait association analysis
GS	Gene TraitSignificance
MM	Module-Membership
IHC	Immunohistochemistry

¹College of Life Sciences, Xinyang Normal University, Xinyang 464000, China. ²College of Medicine, Xinyang Normal University, Xinyang 464000, China. ³Department of Computer Science, City University of Hong Kong, Hong Kong, China. ⁴Department of Recovery Medicine, People's Liberation Army 990 Hospital, Xinyang, China. ⁵Department of Urology, Central Hospital of Xinyang, Xinyang, China. ⁶These authors contributed equally: Lei Wang, Xudong Liu, Zhe Liu and Yafan Wang. ✉email: wangleibio@126.com; luweimed@163.com; xinqiangsong@126.com

TCGA	The cancer genome atlas
GO	Gene ontology
LASSO	Least absolute shrinkage and selection operator
GEPIA	Gene expression profiling interactive analysis
AUC	Area under curve
TIMER	Tumor Immunity Evaluation
OS	Resource overall survival
ROC	Receiver operating characteristic
GSEA	Gene set enrichment analysis

Prostate cancer (PCa), an epithelial malignancy of the prostate, is the second most common cancer in the world and one of the most common malignancies in men^{1,2}. At present, PCa has become a global public problem threatening human health and life³, with more than 1,275,000 new diagnoses and 350,000 deaths each year. The factors associated with the development of PCa are mainly involved in age, race, family heritage, geography, and diet^{4–6}. According to the Epidemiology and End Results report (<https://seer.cancer.gov/>), the five-year survival rate for PCa is 97.5%⁷, and it will account for 13.1% of all new cancers in 2021. Although there are many types of research on PCa, its etiology and pathogenesis remain not fully understood. The current treatments for the PCa include radical surgery⁸, external beam radiotherapy⁷, brachytherapy⁹, local treatment of experimental PCa¹⁰, endocrine therapy¹¹, and chemotherapy¹², etc. Serum prostate-specific antigen (PSA) is now commonly used as the preferred marker for screening PCa, with low specificity¹³. The immune system plays a huge role in preventing and treating of tumors¹⁴. Treatments developed to target the immune system, such as cancer vaccines, can effectively target specific antigens expressed on the surface of cancer cells, thereby protecting normal cells from drug damage¹⁵. However, due to the heterogeneity within tumors and because cancer vaccines target only a limited number of specific antigens, cancer cells that do not express these antigens or a group of cancer cells that have mutations that result in altered surface antigens can evade immune control and thus create new tumor populations that can resist treatment with vaccines encoding the same tumor-associated antigens^{16,17}. Therefore, the development of new effective oncology treatments against heterogeneous populations of tumors has become an urgent task. However, immunotherapy for PCa has not been efficacious for patients in the past¹⁸. Therefore, comprehensive bioinformatics screening for new diagnostic and prognostic indicators and the proposal of new effective cancer drugs are of great importance for the treatment and diagnosis of PCa.

Intratumor heterogeneity is a vital feature of tumorigenesis¹⁹. As cancer cells divide and proliferate, somatic mutations accumulate, some of which give cancer cells a more significant adaptive advantage^{20–23}. Meanwhile, alterations in epigenetics with the tumor microenvironment may lead to different molecular subtypes that have varying tissue types, invasive capabilities, and degrees of differentiation^{24–27}. The presence of heterogeneity makes tumor treatment very challenging, while the discovery of tumor driver genes and associated drugs targeting these genes offers the possibility of treating these tumors^{28,29}. A comprehensive analysis of 333 cases of primary prostate cancer has identified three molecular subtypes³⁰. Precision medicine for PCa needs to consider different molecular subtypes so that tumor heterogeneity can be achieved and patients can be offered personalized treatment options.

Reactive oxygen species (ROS) modulator 1 (ROMO1) is a membrane protein found in mitochondria that is important for regulating mitochondrial ROS production and redox sensing³¹. Romo1 is capable of triggering and exacerbating cancer through extracellular signal-regulated kinases (ERKs) and nuclear factor- κ B (NF- κ B)-induced reactive oxygen species (ROS)³². ROS can trigger and exacerbate cancer in a variety of malignancies^{33,34}. It can also influence cancer cell invasion by affecting the epithelial-mesenchymal transition (EMT) pathway^{35,36}. This protein affects signaling pathways and ROS homeostasis, affects the G2/M phase cell cycle, and leads to cell overgrowth through increased expression levels³⁷. Nevertheless, ROMO1 has not been reported in the tumor microenvironment as well as in the development of prostate cancer.

With the rapid development of next-generation sequencing technologies, bioinformatics analysis allows us to understand tumor characterizes at a multi-omics level³⁸. Research in RNA sequencing (RNA-seq), Co-expression network analysis, immunohistochemistry, immuno-infiltration, and targeted drug studies, etc., are now widely used. RNA sequencing can study the transcriptional profiles of cell populations or the average expression levels of tissues³⁹. The Cancer Genome Atlas (TCGA)⁴⁰, Gene Expression Profiling Interactive Analysis (GEPIA)⁴¹, The Human Protein Atlas (HPA)⁴², etc. are open databases for the collection and storage of patient genomic data. This study explores the PCa RNA-seq data to discover new biomarkers and provide new insights for the diagnosis and treatment of PCa.

Materials and methods

Data acquisition. Gene expression profiles and clinical data for PCa were downloaded using the UCSC Xena (<http://xena.ucsc.edu/>) database⁴³, containing 436 PCa tissue and 114 normal tissues. Patients with unclear survival time and survival status characteristics were excluded from the cohort, and all data were collected on 10 April 2021.

Determination of cluster count and membership. Based on the expression profile of PCa, Consensus-ClusterPlus (v. 1.52.0) R language software package has carried out unsupervised clustering classification for 550 samples^{44,45}. The project selection criteria were selected 1000 iterations and 80% resampling rate Pearson correlation. Concurrently, principal component analysis for further filtering of the results of consensus clustering is employed. Reshape2 (v. 1.4.4) and factoextra (v. 1.0.7) two R language software packages were used to reduce the dimensions of clusters of different clusters to verify the results, and the first two principal components of

the principal component analysis were selected for subsequent analysis. Also, all samples are within the cluster and within the cutoff height of ($-100 \sim 150$) and cutoff width of ($-200 \sim 150$). The ggplot2 (v. 3.3.3) R language software package was used to draw the graph to display the results.

Identification of differentially expressed genes based on consensus clustering and PCA. The Limma (v. 3.46.0) R language package was used to screen differentially expressed genes (DEGs) among non-cancerous samples with different prostate carcinoma subtypes⁴⁶. The DEGs are identified by significant cutoffs. Genes with $|\log_2FC|$ (fold change) ≥ 1 and P -value ≤ 0.01 is considered as DEGs. The R language software packages of ggplot2 (v. 3.3.3) and RColorBrewer (v. 1.1.2) were used to draw the graph to display the results. The Venn diagram of the overlapping DEGs was outputted by the Draw Venn diagram website (<http://bioinformatics.psb.ugent.be/webtools/Venn/>), and the Upset plot was drawn by the UpSetR R language package.

Functional enrichment analysis of noncancerous samples and different tumor subtypes of DEGs. To elucidate the biological functions of the DEGs, we performed further functional annotation and pathway enrichment analysis using the Metascape (<https://metascape.org>) online tools⁴⁷. P -value < 0.05 was considered as the cutoff value. An R package heatmap is used in heatmap figure. And two-way hierarchical clustering is used for clustering.

Gene co-expression network analysis to obtain key gene modules associated with tumorigenesis and malignant progression. Gene co-expression network analysis is widely used in identifying functional gene modules associated with human disease and has proven to be highly effective^{48,49}. The complexity of tumors is further expressed through intratumoral heterogeneity and the tumor microenvironment, which encompasses intratumoral cell differences and interactions and the expression of several biomarkers between normal and different tumor cells^{50,51}. In this study, we used Weighted Gene Co-expression Network Analysis (WGCNA) and Multiscale Embedded Gene Co-expression Network Analysis (MEGENA) to identify key genes associated with prostate carcinogenesis as well as tumor malignant progression WGCNA (v. 1.70-3) R language package was used to perform a WGCNA of all DEGs with clinical traits, mining modules, and key genes associated with clinical characteristics⁵²:

1. The gene correlation matrix was transformed into a scale-free network.
2. Unsupervised hierarchical clustering was performed using a dynamic tree chopping algorithm, and the clustered tree branches were defined as modules.
3. The correlation between each module and clinical traits was calculated, and the modules with higher correlations were selected.

Subsequently, the hub genes in the selected modules were ranked by intra-module connectivity and correlation with module trait genes, and finally, candidate genes were identified.

MEGENA is an innovative co-expression network analysis method that offers unique advantages over WGCNA for efficiently constructing large-scale co-expression planar filter networks and preserving gene interactions⁵³. The R software package MEGENA (v. 1.3.7) is used to perform MEGENA, which consists of the following steps: (1) constructing a planar filtered network (PFN); firstly, calculating correlation coefficients based on gene expression profiles, and then filtering and clustering gene pairs using a parallel filtering method to obtain a fast planar filter network; (2) multi-scale clustering analysis; from the initial PFN of the connected components, multi-scale clustering of each parent cluster can obtain more sub-modules, followed by hierarchical clustering results; (3) downstream analysis, using multiscale hub analysis (MHA) to identify important hubs based on the network topology; (4) Finally, the correlation between clustering results and clinical information was analyzed by cluster-trait association analysis (CTA). To test whether selected modules in WGCN and multiscale hub genes in MEGCN were highly associated with tumorigenesis and malignant progression, we performed GO and KEGG pathway⁵⁴ analysis on all genes in the selected DAVID (<https://david.ncifcrf.gov/>) database⁵⁵ for hub modules. $P \leq 0.05$ was considered as the cut-off criterion for identifying enrichment. Similarly, an R package heatmap is used in heatmap figure. And two-way hierarchical clustering is used for clustering.

Validation and survival analysis of pivotal genes. In WGCNA, when hub modules were identified, we selected candidate hub genes by module connectivity ($\text{cor. gene module-Membership (MM)} > 0.8$) and clinical feature correlation ($\text{cor. gene TraitSignificance (GS)} > 0.2$), both by Pearson correlation absolute values were used to determine. At the same time, we chose the STRING database (<https://www.string-db.org/>)⁵⁶ to identify the connectivity of candidate genes. Next, Cytoscape's plug-in CytoHubba was used to identify top50 genes as candidate genes⁵⁷. The gene modules from the WGCNA and MEGENA analyses are saved separately. The userListEnrichment R function was used to calculate the overlap of genes between the WGCNA and MEGENA modules. The WGCNA and MEGENA modules that overlapped significantly and were significantly associated with PCa were retained. In this study, we used the Gene Expression Profiling Interactive Analysis (GEPIA) dataset (<http://gepia.cancer-pku.cn/>) for prognostic analysis of PCa⁵⁸. We analyzed the differential expression of key hub genes in PCa and their association with RFS using GEPIA. P -value and fold change were defined as 0.05 and group limit of 50% for two survival analyses. Also, normal distribution tests and differential analysis of hub genes in normal versus tumor tissue and normal versus different tumor subtypes were performed based on the Wilcoxon test. Finally, Human Protein Atlas Dataset (<https://www.proteinatlas.org/>)⁵⁹ was used to validate the immunohistochemistry (IHC) of the hub gene.

Development and validation of the hub gene prognostic model. To assess the prognostic value of hub genes, Cox regression analysis was used to evaluate the correlation between genes and survival status in a cohort of 494 PCas. Next, we chose the glmnet R language package for LASSO Cox regression modeling to narrow down the candidate genes and build prognostic models⁶⁰. Ultimately, the screening retained three genes and their coefficients, and the penalty parameter λ was determined based on the minimum criterion. The algorithm obtained a more accurate model by constructing a penalty function. As a method for complex covariance data, variable selection can be achieved and parameter estimation to better address multiple covariances in regression analysis. The PCa expression data were centrally normalized using a scale function, and a risk score was calculated.

Patients with PCa OS were divided into low and high-risk subgroups based on median risk scores. Kaplan Meier analysis was performed to compare OS time assessment risk scores and overall survival between the two subgroups. To validate the accuracy and predictive power of the model, ROC curve analyses were performed using the *survivor*, *survminer*, and *timeROC* R packages for one year, three years, and five years, respectively, to calculate the area under the curve (AUC) and to compare the effect of the *timeROC* R package on classifier performance.

Relationship between key genes and tumor-infiltrating immune cells. The Tumor Immunity Evaluation Resource (TIMER) (<https://cistrome.shinyapps.io/timer/>) is a comprehensive resource for detecting immune cell infiltration in tumor tissue using RNA-Seq expression profiling data and contains immune infiltration in different cancer types⁶¹. Hub gene expression was correlated with six types of immune infiltration (B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, and dendritic cells) were correlated and assessed using the Gene module. The SCNA module was also used to explore the correlation between somatic cell copy number changes and the abundance of immune infiltrates.

Exploring the drug sensitivity of the hub gene. The GSCALite (<http://bioinfo.life.hust.edu.cn/web/GSCALite/>) database is a genomic cancer analysis platform⁶². The database can be used for genomic and immunogenomic analyses. It also enables researchers to combine clinical information and small molecule drugs to mine candidate biomarkers and valuable small molecule drugs for better experimental design and further clinical trials. The GSCA data contains 33 cancer types from TCGA and normal tissue data from GTEx, and over 750 small-molecule drugs from GDSC and CTRP for 10,000 genomic data. Spearman correlations represent gene expression associated with drugs. A positive correlation means that a gene with high expression is resistant to the drug and vice versa.

Statistical analysis. R⁶³ (version 3.6.2) and related software packages were an application for all the statistical analyses. $P < 0.05$ is a statistically significant difference.

Result

Combined consistency clustering and principal component analysis to obtain sample cohorts. The analytical process used in this study is shown in Fig. S1. To follow up on the molecular heterogeneity of PCa, we performed an unsupervised consensus analysis on all samples. In this study, we chose $k = 5$, which allowed us to divide all samples into five groups (Fig. S2). Among them, C1:59, C2:146, C3:129, C4:102, and C5:114. Moreover, we found that C1-C4 were tumor subgroups, and C5 were normal samples (Fig. 1A-C). Next, to verify the robustness of this classification, we performed another principal component analysis based on the results of consistent clustering and observed the subgroup differentiation. PC1 and PC2 were selected as the main components for the analysis, and 260 samples with significant differences were obtained after excluding outlier samples as well as those with insignificant differences (Table 1). As shown in Fig. 1D, there was an individual crossover between tumor subgroups, indicating good differentiation between subgroups. The clustering results were as follows: C1:41, C2:68, C3:72, C4:51, and C5:28, reflecting the impact of differences between tumor subgroups and normal samples on transcriptional profiles.

Identification of differentially expressed genes and GO functional annotation in PCa. We performed differential gene expression analysis of the four tumor subgroups separately from normal samples. A total of 3521 differentially expressed genes were obtained. Of these, 581 up-regulated and 1432 down-regulated genes were obtained for C1 vs. C5, 231 up-regulated and 637 down-regulated genes were obtained for C2 vs. C5, 234 up-regulated and 684 down-regulated genes were obtained for C3 vs. C5, and 282 up-regulated and 2265 down-regulated genes were obtained for C4 vs. C5 (Fig. 2A). Comparisons between tumor subgroups revealed 842 DEGs specific to C1, 47 DEGs specific to C2, 29 DEGs specific to C3, and 1045 DEGs specific to C4 (Fig. 2B). Three hundred ninety-six overlap genes were also identified, suggesting a degree of similarity in expression profiles between tumor subgroups but also a high degree of heterogeneity between tumor subgroups (Fig. 2C).

Functional enrichment analysis of four groups of differentially expressed genes. The researchers used the Metascape website to carry out a biological process analysis of the GO function of 3521 DEGs (Fig. S3). As shown in Fig. 3A, there was a strong interaction between the four groups of differentially expressed genes. Also, DEGs were found to be mainly enriched in functions such as leukocyte migration, inflammatory response, regulation of MAPK cascade, circulatory system processes cell component morphogenesis, and regulation of cell adhesion (Table 2). Moreover, Metascape enrichment analysis also showed that DEGs differed significantly between the four groups in the pathways regulating cell adhesion and leukocyte migration, demon-

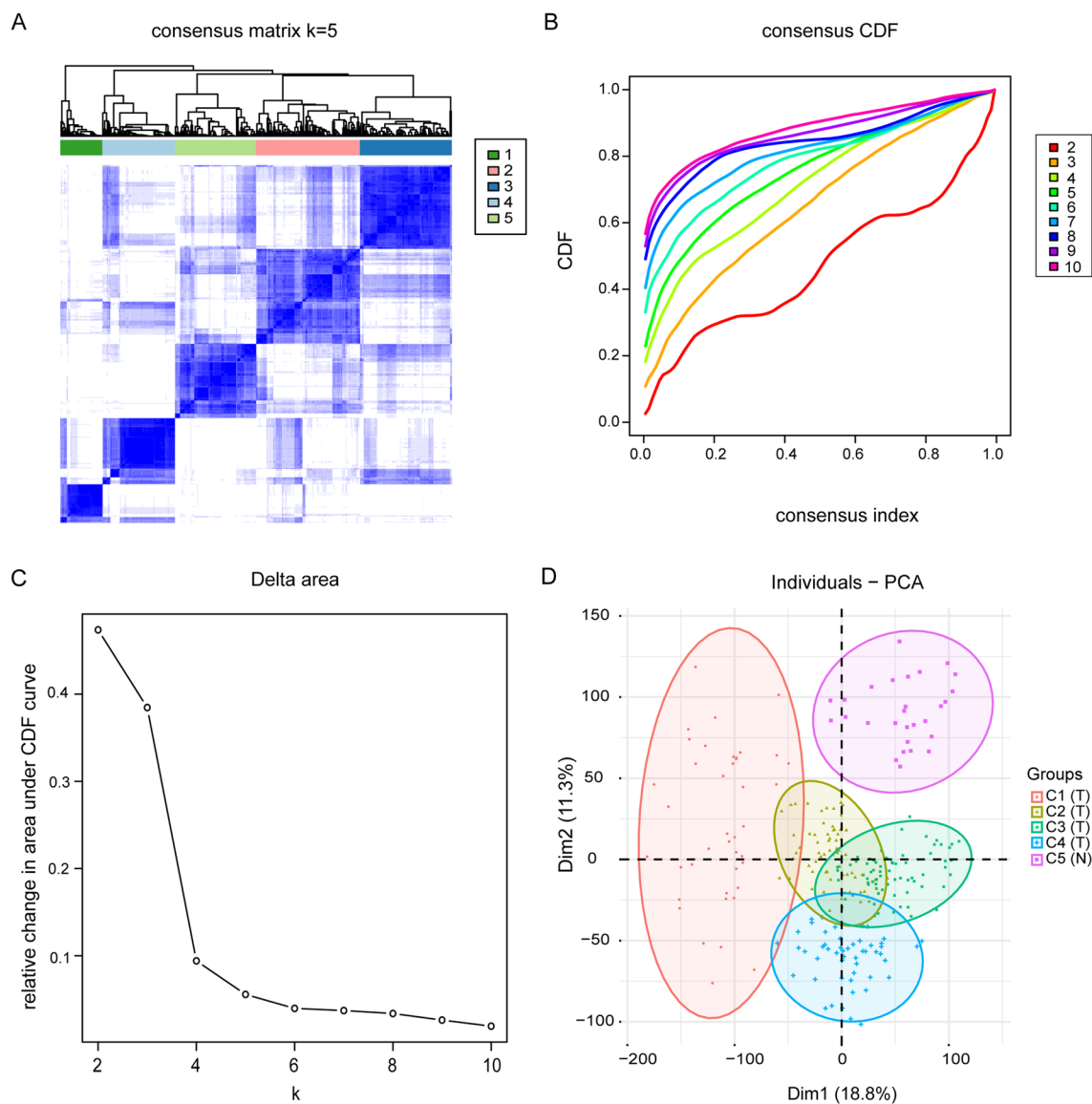


Figure 1. PCA tumor subtypes classification and identification in TCGA cohort. **(A)** The 550 samples were split into five clusters by the consensus clustering matrix ($k=5$). **(B)** The consensus clustering approach established the cumulative distribution function (CDF) curve of Top 10. **(C)** The CDF Delta area curve of all samples with different k values. **(D)** Principal component analysis was carried out based on the results of consensus clustering.

strating genetic and functional differences in tumor subtypes (Fig. 3B–E). Naturally, to ensure that the differential analysis did not filter out key genes, we also performed GSEA analysis on all genes and found that they were mainly enriched in P53, DNA repair, Notch signaling pathway, etc. (Fig. S4).

Co-expression network analysis reveals key modules of PCa network. We used the expression profiles and clinical information of 260 samples from PCa as input files for WGCNA by integrating them. We eliminated two outlier samples based on the clustering tree and then plotted the sample dendrograms and expression heat maps of the traits (Fig. 4A). After excluding the two outlier samples, the remaining 258 samples and 3521 differentially expressed genes were subjected to WGCNA analysis (Fig. S5). We chose a soft threshold power of 8 (scale-free $R^2=0.85$) by calculation, ensuring that the scale-free network was reasonable (Fig. 4B). We built the network by choosing the soft threshold. And the minimum number of genes in the modules was required to be 50. Subsequently, the initial modules were divided using a dynamic tree, and the genes were divided into modules based on the similarity of the characteristic genes in the gene clustering map, and ten gene modules were finally identified (Fig. 4C). One color indicates a gene module, while grey gene modules are considered invalid genes that cannot be assigned to a module. Next, we created a heat map of the correlation between gene modules and clinical information (Fig. 4D). We found that the blue module was found to have the highest correlation coefficient with consensus clustering (Pearson $cor=0.87$, $p=1e-200$) (Fig. 4E). This suggests

Variable	Type	%
PCA (260)		
Tumor	232	89
Normol	28	11
CCF		
Cluster1	41	16
Cluster2	68	26
Cluster3	72	28
Cluster4	51	20
Cluster5	28	10
Relaplse		
NA	13	5
NO	206	79
YES	41	16
Age		
> =60	163	63
<60	97	37
Gleason_score		
6	17	6
7	144	55
8	28	11
9	69	27
10	2	1
OS		
0 < = OS < = 1000	142	55
1000 < OS < = 2000	84	32
2000 < OS < = 3000	24	9
3000 < OS < = 4000	5	2
4000 < OS < = 5000	3	1
OS > = 5000	2	1

Table 1. Information on the clinical features of patients with prostate cancer.

that the genes in this module correlate with tumor typing as well as malignancy; meanwhile, we found a large correlation coefficient between the green module and tumorigenesis (Pearson $\text{cor}=0.57$, $p=3e-08$) (Fig. 4F). In contrast, the absolute values of the correlation coefficients between genes and age, recurrence status, and gleason_score were smaller in the other modules, suggesting that the genes are less relevant to other clinical information. Finally, we calculated the connectivity of the genes within the blue (Fig. 5A) and green modules (Fig. 5B) and selected the top50 most closely connected genes for subsequent analysis.

Meanwhile, the expression profiles of 3521 differentially expressed genes (DEGs) were applied to the multi-scale embedded gene co-expression network (MEGENA) analysis. After parallelization, early termination, and pre-quality control steps, 10,333 gene pairs of PFNs were used to enter 1755 MCA for multiscale cluster identification. A total of 5 different scales were identified by multi-scale cluster analysis, and 226 significant modular clusters were constructed (Fig. S6), where the modular p-values were all less than 0.05.

Comparing the significant modules of WGCNA with the MEGENA modules, we found a degree of overlap between the key gene modules of WGCNA and those of MEGENA. We found a large degree of overlap between the blue module and green module of WGCNA and the C1_5 module (Fig. 6A) and C1_7 (Fig. 6B) module of MEGENA. We also annotated the genes of the four modules and found that the four groups of genes are mainly involved in the cell cycle, cell division, and transcriptional functions (Fig. S7). WGCNA and MEGENA can complement each other and help us to preserve the maximum range of gene modules specific to PCa. We found that the expression profiles of this class of overlapping genes differed greatly between tumor and normal adjacent tissue, and some of the genes also differed greatly between different subtypes of tumors (Fig. 7A, B). Further, the expression of NDUFB7, AURKAIP1, ROMO1, SCAND1, GADD45GIP1, and NDUFA13 was upregulated in the four tumor subtypes compared with normal adjacent tissue (Fig. 7C). Next, we annotated the GO_BP function of 29 overlapping genes based on the DAVID database (Table 3) and found that these genes are mainly involved in cell division, cell cycle, mitosis, and positive regulation of ubiquitin-protein ligase activity (Fig. 7D).

Validation of key candidate genes. We first performed a disease-free survival analysis of key genes using the GEPIA database to explore the relationship between genes and phenotype and prognostic status. We selected median as a criterion to differentiate gene expression levels and found that TPX2, ROMO1, PLK1, UBE2C, and KIF4A were considered hub genes with the most significant p-values selected from the 29 core genes (Fig. 8A).

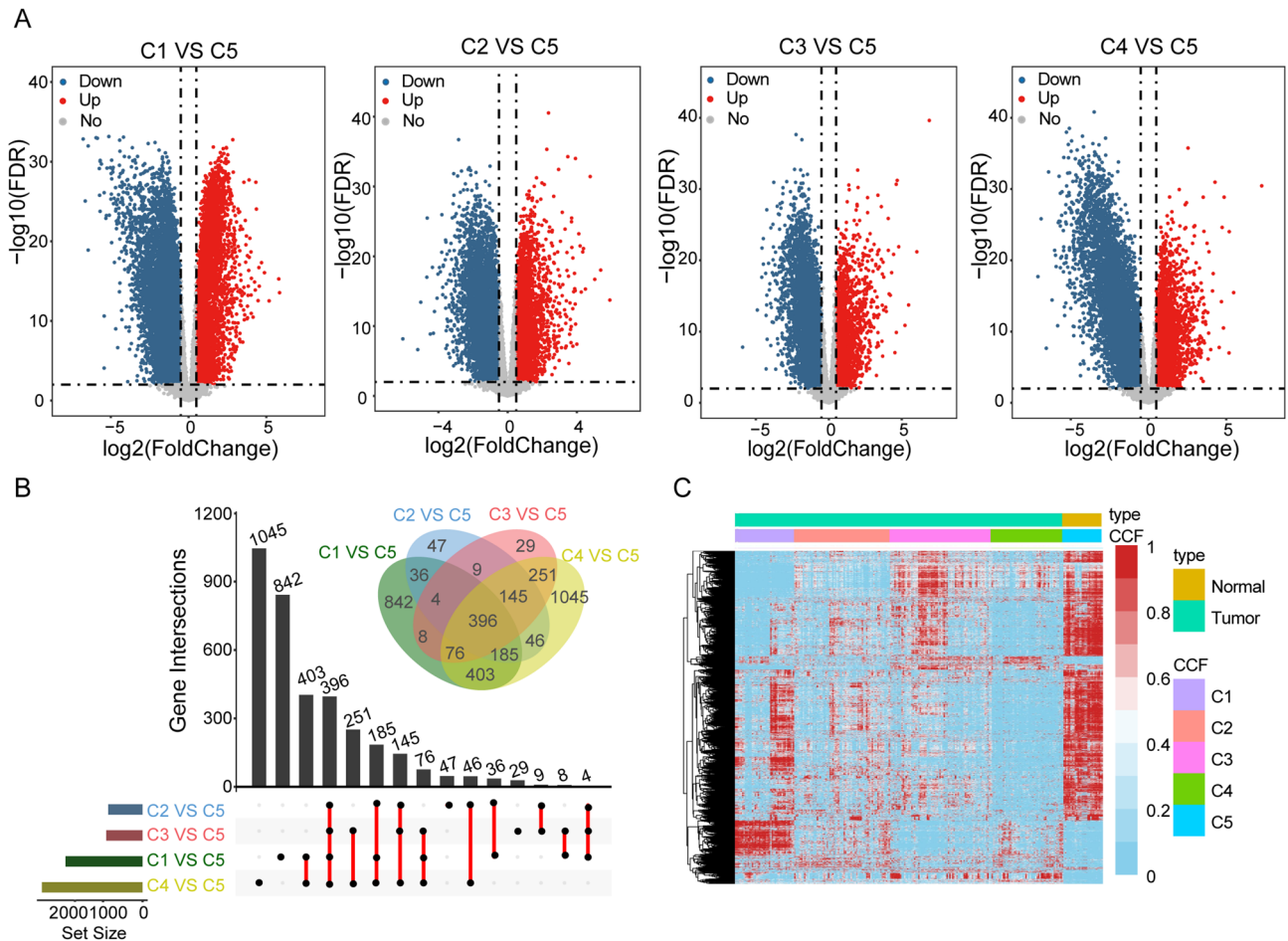


Figure 2. Analysis of differences between four tumor subtypes and normal samples. (A) Volcano plot of four groups, with $|\log_2(\text{FC})| > 2$ and $\text{FDR} < 0.01$. (B) Venn diagram and Upset plot were used to visualize common genes between 4 clusters datasets. The number of genes annotated is presented on the y-axis. (C) The heat map shows the expression profiles of differentially expressed genes.

We then perform a Wilcoxon test on the five hub genes screened, and we found significant differences ($P < 0.05$) in the expression of the five genes between tumor and normal adjacent tissue (Fig. 8B). Meanwhile, we further analyzed the expression differences of the five hub genes among different tumor subtypes, and we found that these genes also differed significantly among tumor subtypes of PCa (Fig. 8C). Moreover, we found that ROMO1 was highly expressed in both tumor and normal tissues after immunohistochemistry (Fig. 8D). Meanwhile, we incorporated PCa data from GETx that found significant upregulation of ROMO1 in tumor tissues from GEPIA database (Fig. S8).

Construction and validation of prognostic gene markers in TCGA. The data for 497 tumors were obtained from The Cancer Genome Atlas (TCGA) dataset, and corresponding clinical information was incorporated into the prognostic training model. The expression data of five genes, TPX2, ROMO1, PLK1, UBE2C, and KIF4A, were pooled, and LASSO Cox was applied to construct a regression model. The model identified ROMO1, PLK1, and KIF4A5 based on the optimal value of λ to construct an OS prognostic model for PCa patients (Fig. 9A, B). The risk scores, patient survival status distributions, and gene expression profiles associated with the three genetic features in the training dataset are shown in Fig. 9C. OS was significantly lower in the high-risk group than in the low-risk group ($P = 0.0021$) (Fig. 9D). Finally, ROC analysis was performed, and the area under the ROC curve (AUC) for overall survival (OS) at 1, 3, and 5 years was 0.789 and 0.72 and 0.626, respectively (Fig. 9E). In a subsequent study, we also did a LASSO regression analysis on ROMO1 and found that ROMO1 could also be used as an independent prognostic model and that its low-risk group had a better survival outcome than the high-risk group (Fig. S9).

Alterations in the immune microenvironment may result from aberrant regulation of key genes in PCa. To explore the relationship between hub genes expression and tumor-infiltrating immune cells, we used the TIMER tool to analyze the correlation between key gene expression and the level of immune infiltration. Our results showed that ROMO1 expression was strongly correlated with immunity to PCa and was found to be negatively correlated with B cells, CD8 + T cells, macrophages, neutrophils, dendritic cells and posi-

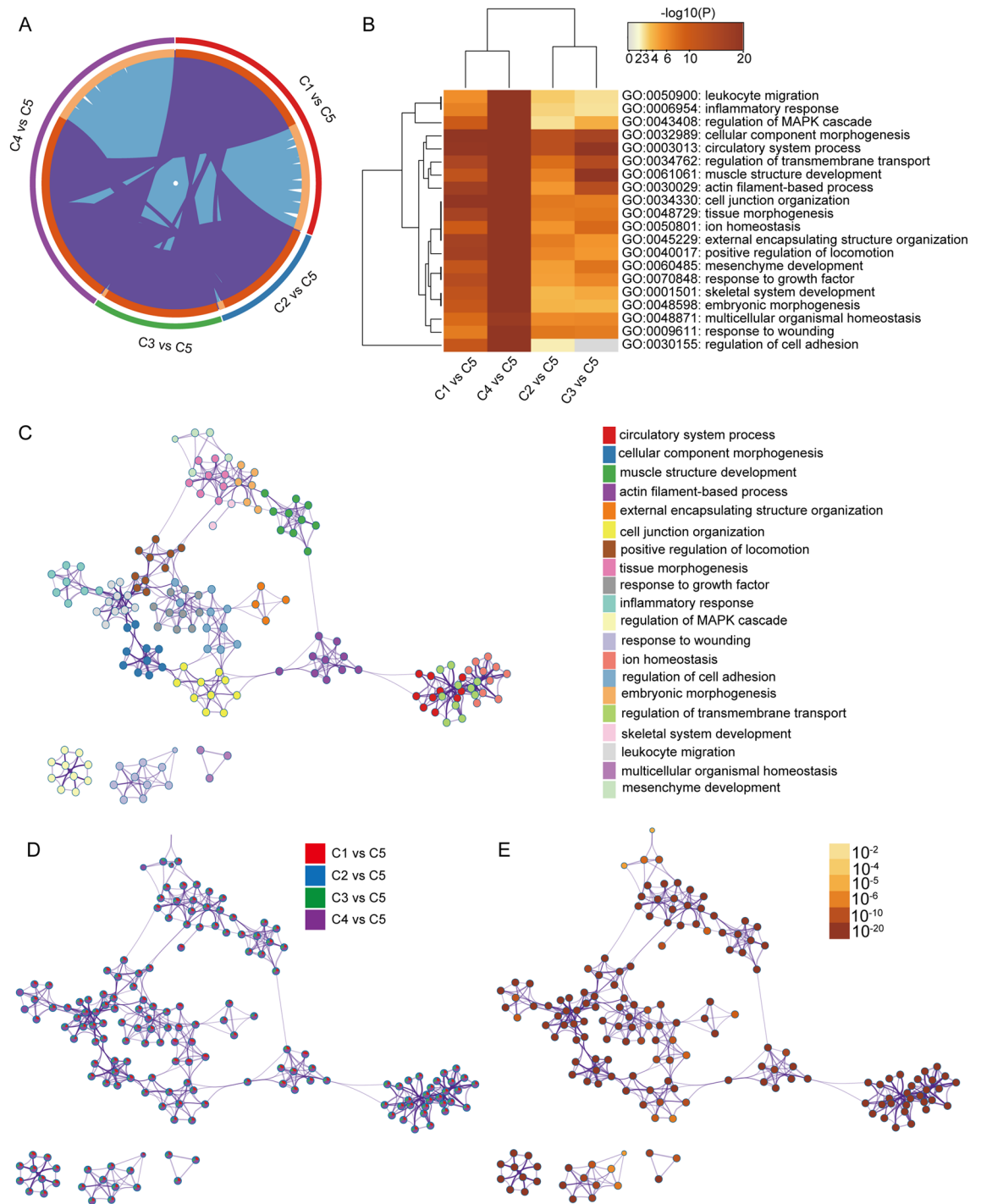


Figure 3. Meta-enrichment analysis summary for PFNs lists in 4 comparison cohorts. **(A)** The overlap of differentially expressed genes in the four selected groups is shown in the circos plot explored using Metascape. **(B)** Each box indicates whether the gene in each gene list (column, C1 vs C5, C4 vs C5, C2 vs C5, and C3 vs C5) enriched in each selected top 20 GO term. The darker the color, the more significance the *P*-value. R package heatmap is used in this process, and two-way hierarchical clustering is used for clustering. **(C)** The network of enriched terms is colored by cluster-ID, and nodes that share the same cluster are typically close to each other. **(D)** Network enrichment terms and genes are colored by the database, where the terms containing more genes tends to have more significates. **(E)** Coloring by the degree of enrichment, with darker colors indicating a greater number of genes enriched to that pathway or biological process class.

tively correlated with CD4+ T cells; PLK1 and KLF4A expression was positively correlated with B cells, CD8+ T cells, macrophages, neutrophils, dendritic cells (Fig. 10 A-C). In addition, we found that different mutant forms

No	Category	Term ID	Description	LogP	Ratio
1	GO_BP	GO:0003013	Circulatory system process	-71.68	207/611
2	GO_BP	GO:0032989	Cellular component morphogenesis	-65.57	226/766
3	GO_BP	GO:0061061	Muscle structure development	-56.30	186/609
4	GO_BP	GO:0030029	Actin filament-based process	-55.21	215/794
5	GO_BP	GO:0045229	External encapsulating structure organization	-52.90	143/398
6	GO_BP	GO:0034330	Cell junction organization	-51.96	195/701
7	GO_BP	GO:0040017	Positive regulation of locomotion	-48.52	171/587
8	GO_BP	GO:0048729	Tissue morphogenesis	-46.92	177/638
9	GO_BP	GO:0070848	Response to growth factor	-46.55	192/736
10	GO_BP	GO:0006954	Inflammatory response	-45.06	196/778
11	GO_BP	GO:0043408	Regulation of MAPK cascade	-44.56	185/712
12	GO_BP	GO:0009611	Response to wounding	-41.77	172/658
13	GO_BP	GO:0050801	Ion homeostasis	-40.81	190/787
14	GO_BP	GO:0030155	Regulation of cell adhesion	-40.19	181/734
15	GO_BP	GO:0048598	Embryonic morphogenesis	-39.35	154/569
16	GO_BP	GO:0034762	Regulation of transmembrane transport	-38.50	152/565
17	GO_BP	GO:0001501	Skeletal system development	-37.11	137/486
18	GO_BP	GO:0050900	Leukocyte migration	-36.41	140/511
19	GO_BP	GO:0048871	Multicellular organismal homeostasis	-35.28	139/516
20	GO_BP	GO:0060485	Mesenchyme development	-34.80	99/288

Table 2. Functional annotation of GO_BP for top20 of four sets of DEGs.

of ROMO1 correlated with immune infiltration of CD8 + T cells, macrophages, and neutrophils; different mutant forms of KLF4A correlated with immune levels of CD8 + T cells, macrophages, neutrophils, and dendritic cells; and we found that ROMO1, PLK1, and KLF4A were found to have higher levels of immune infiltration in dendritic cells (Fig. 10 D–F). This also reveals the influence of key genes on the immune microenvironment of PCa.

Drug sensitivity and resistance of key genes. To investigate the relationship between the expression of hub genes and drug interactions, we used the GSCA database to screen for drugs that interacted with it. The GSCA database integrated drug sensitivity and gene expression profile data for cancer cell lines from both the GDSC and CTRP databases into GSCALite, and then, using Spearman correlation analysis, the expression levels of genes were correlated with the IC50. FDR < 0.05 was considered significant (Supplementary Table 1). During the research, we found that ROMO1 expression was significantly positively correlated with nutlin-3, fluorouracil, etc. KLF4A expression was significantly positively correlated with PD318088 and selumetinib, and negatively correlated with ML239; PLK1 expression was significantly negatively correlated with STF-31 was significantly negatively correlated (Fig. 11). This reveals a diversity of expression levels of key genes and drug sensitivity.

Discussion

In recent years, PCa has become the most common and deadly solid cancer and genitourinary tumor in men worldwide⁶⁴, with a diagnosis rate of 12% and a mortality rate of 9%, and its incidence rises with age⁶⁵. The traditional clinical treatment options include resection, radiotherapy, chemotherapy, and endocrine therapy⁶⁶. The tumor heterogeneity leads to limitations in conventional treatment and makes it difficult to manage and risk assess patients effectively^{67,68}. Recent studies have demonstrated the imperative of identifying new effective biomarkers and promising immune-related therapeutic targets that can be used to guide cancer treatment in clinical practice⁶⁹. In this study, we found a high correlation between ROMO1 expression and clinical features such as the occurrence of PCa, subtypes classified by consistent clustering. Also, ROMO1 correlated with the level of immune cell infiltration and immune pathways in PCa.

Firstly, many factors can contribute to cancer development, such as dietary status, genetic mutations, epigenetic alterations, etc. We consequently identified 260 valid samples comprising a normal group and four tumor sub-types by consensus clustering and principal component analysis after outliers were excluded. Secondly, the limmaR package was selected to differentially analyze the four groups of tumor subtypes from the normal group, respectively. Finally, 3521 DEGs were identified and functionally annotated, and these differentially expressed genes were found to be mainly enriched in an immune-related pathways. Next, WGCNA and MEGENA were applied to construct co-expression networks of differentially expressed genes to identify gene modules associated with tumorigenesis and heterogeneity, and 29 overlapping genes were identified by calculation. Then, the Lasso Cox regression model was constructed to identify ROMO1, PLK1, and KIF4A5 as the optimal core genes. Subsequently, when analyzing the relationship between the expression of hub genes and tumor-infiltrating immune cells, it was found that the expression of ROMO1, PLK1, and KIF4A were all associated with tumor-infiltrating immune cells, which in turn led to the alteration of the tumor microenvironment alterations, which in turn lead to tumor heterogeneity. Finally, we found that nutlin-3, fluorouracil, and others could be potential therapeutic

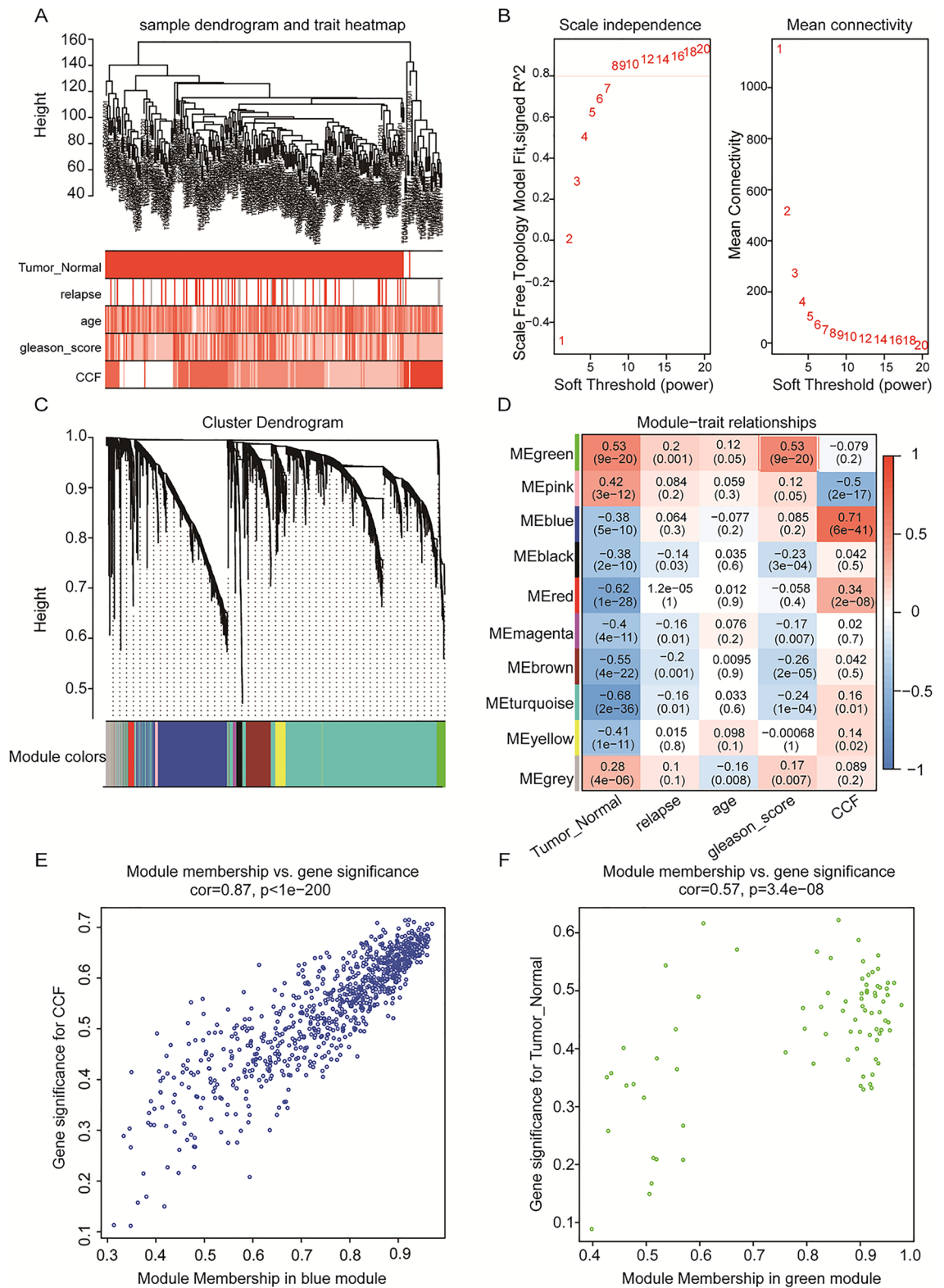


Figure 4. Analysis of weighted gene co-expression networks for the construction of PCA. **(A)** Dendrogram combining clinical data and transcriptional expression profiles for sample Euclidean distance clustering. **(B)** Analysis of soft threshold power for WGCNA: (left) Analysis of the scale-free matching index (β) for various soft threshold powers. (Right) Analysis of the average connectivity of various soft threshold powers. **(C)** Dendrogram of all differentially expressed genes clustered based on the measure of variability (1-TOM). The colors bands show the results obtained from the automated individual cluster analysis. **(D)** The heat map shows the relationship between MEs and clinical characteristics. Each cell contains the corresponding correlation coefficient and p-value. **(E, F)** Scatterplot of gene significance versus module affiliation in the three selected key modules blue, green modules.

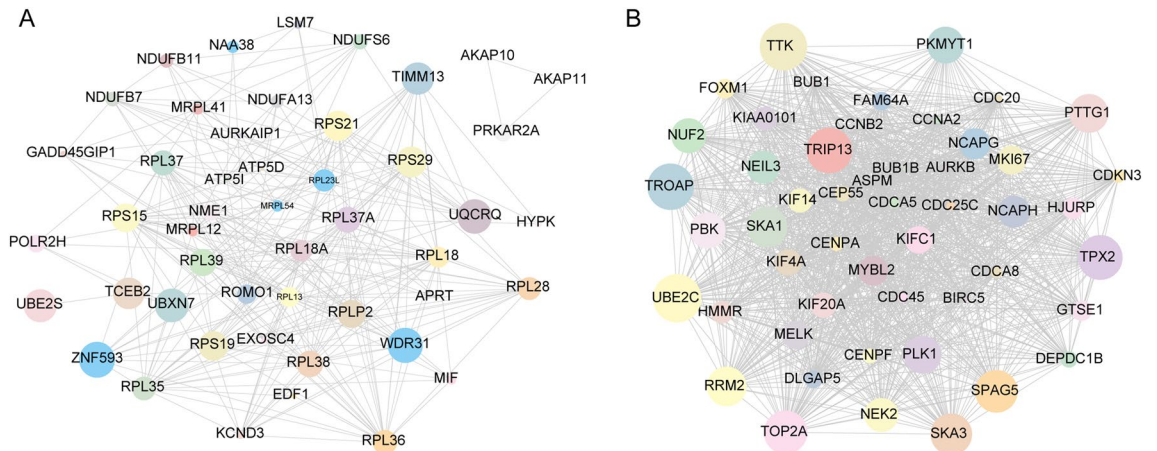


Figure 5. Weighted gene co-expression network analysis of gene interaction networks in key gene modules of PCa (A). The top 50 genes with the highest levels of intramodular connectivity in the blue module. (B). The top 50 genes with the highest levels of intramodular connectivity in the green module. (The size of the circle represents combined scores).

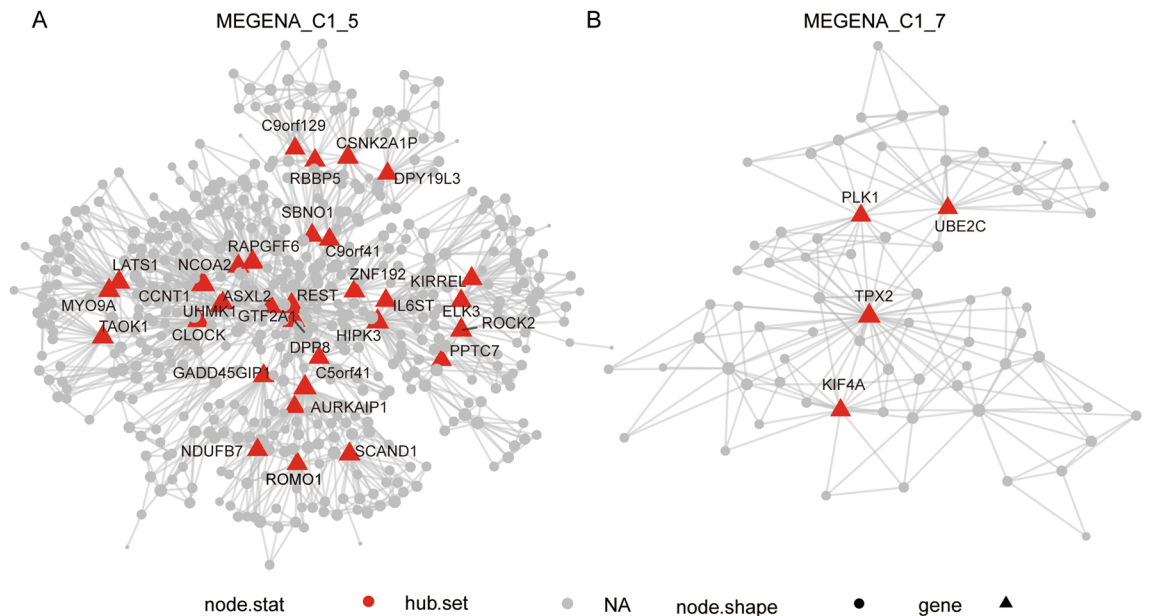


Figure 6. Visualization of interest modules and hub genes multiscale embedded gene co-expression network analysis (MEGENA). (A). Intergenic connectivity of genes in MEGENA_C1_5 and identification of key genes. (B). Intergenic connectivity of genes in MEGENA_C1_7 and identification of key genes. The red triangle represents the HUB gene and the gray circle represents the non-core gene.

agents for ROMO1, and PD318088 and selumetinib could be potential therapeutic agents for PLK1. In conclusion, our study provides a new theoretical basis for the diagnosis and treatment of PCa.

Polo-like kinase 1 (PLK1) is a member of a family of serine/threonine protein kinases that are widely found in eukaryotic cells⁷⁰. Its specific functions are mainly in cell cycle processes, including controlling mitotic entry and the G2/M checkpoint, coordinating centrosomes and the cell cycle, regulating spindle assembly and chromosome segregation, performing multiple functions during mid-spindle and abscission, promoting DNA replication, participating in the cytoplasmic division and meiosis, and playing an important role in the initiation, maintenance, and completion of mitosis^{71–73}. The pharmacological inhibition of PLK1 in triple-negative breast cancer has been reported to increase the anti-proliferative activity of drug-resistant cells, which in turn causes G2/M phase block and increases the phosphorylation of cell cycle proteins inducing apoptosis⁷⁴. Also, in PCa, mitotic kinase polo-like kinase 1 (PLK1) is expressed at elevated levels and is associated with tumor grade⁷⁵. In the present study, we also found a significant prognostic effect of PLK1, with expression in different subtypes of the prostate gland also differing significantly from normal samples, and found that PLK1 expression was also associated with tumor-infiltrating immune cells, further demonstrating the reliability of PLK1 as a biomarker for PCa.

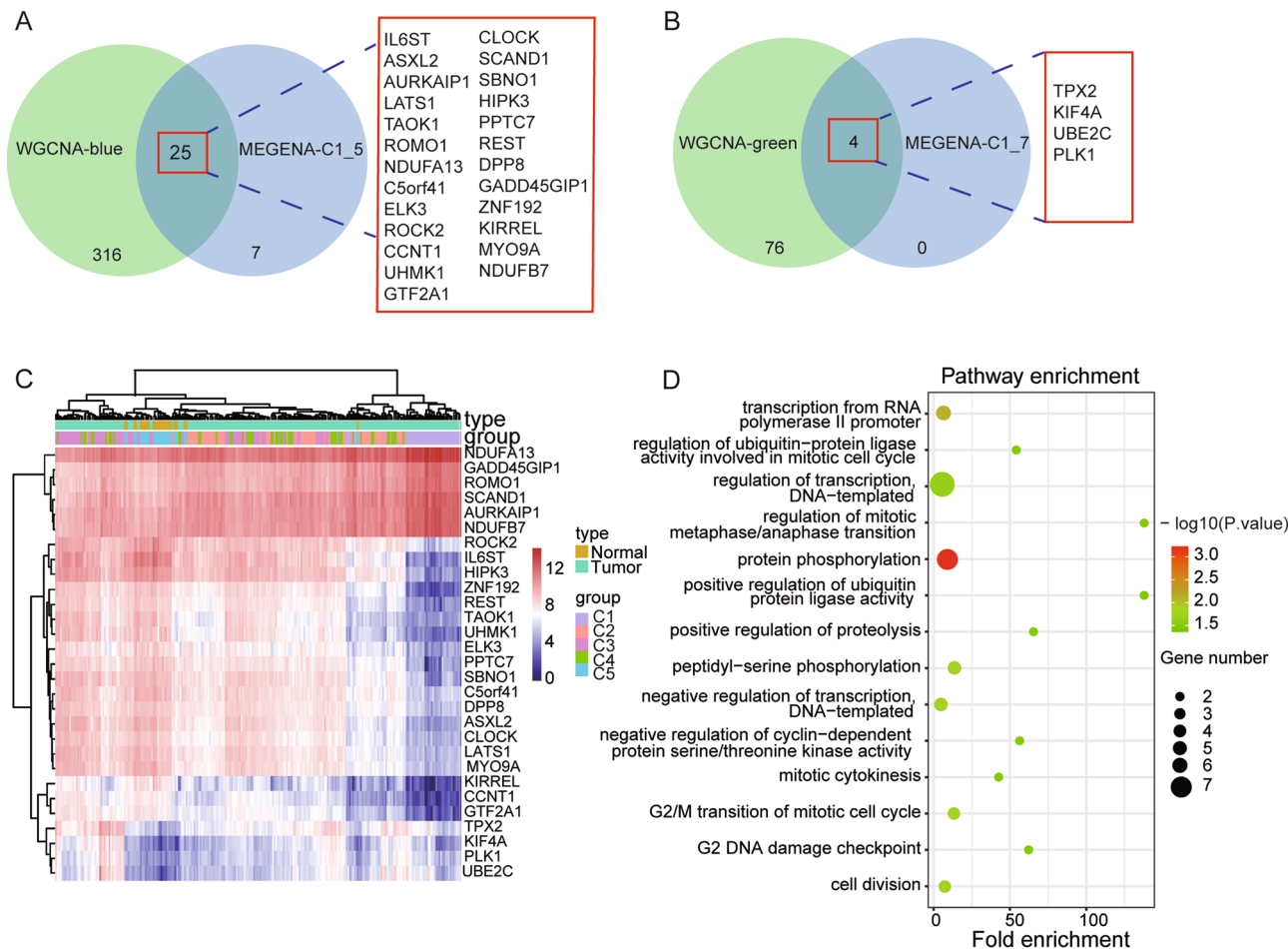


Figure 7. Visualization and functional annotation of candidate genes. Venn diagram showing all identified hub genes and their overlap in WGCNA-blue and MEGENA-C1_5 (A), and WGCNA-green and MEGENA-C1_7 (B). (C). Each box indicates the expression level of each candidate gene (row) in each sample (column). The red indicates the high expression, while the blue indicates the low expression. Similarly, an R package heatmap is used in this figure. And two-way hierarchical clustering is used for clustering. (D). GO enrichment analysis was performed on the candidate genes.

No	Term category_BP	Count	P value	Fold.enrichment	FDR
1	GO:0006468:protein phosphorylation	6	0.00	8.18	0.14
2	GO:0006366:transcription from RNA polymerase II promoter	5	0.00	6.06	0.64
3	GO:1904668:positive regulation of ubiquitin protein ligase activity	2	0.01	138.21	0.62
4	GO:0030071:regulation of mitotic metaphase/anaphase transition	2	0.01	138.21	0.64
5	GO:0018105:peptidyl-serine phosphorylation	3	0.02	14.93	0.64
6	GO:0,051,301:cell division	4	0.02	7.11	0.64
7	GO:0000086:G2/M transition of mitotic cell cycle	3	0.02	13.62	0.64
8	GO:0006355:regulation of transcription, DNA-templated	7	0.02	2.89	0.69
9	GO:0045862:positive regulation of proteolysis	2	0.02	65.47	0.69
10	GO:0031572:G2 DNA damage checkpoint	2	0.03	62.19	0.69
11	GO:0045736:negative regulation of cyclin-dependent protein serine/threonine kinase activity	2	0.03	56.54	0.69
12	GO:0051439:regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle	2	0.05	54.08	0.69
13	GO:0045892:negative regulation of transcription, DNA-templated	4	0.04	4.99	0.74
14	GO:0000281:mitotic cytokinesis	2	0.04	42.89	0.74

Table 3. Functional annotation of GO_BP for candidate genes.

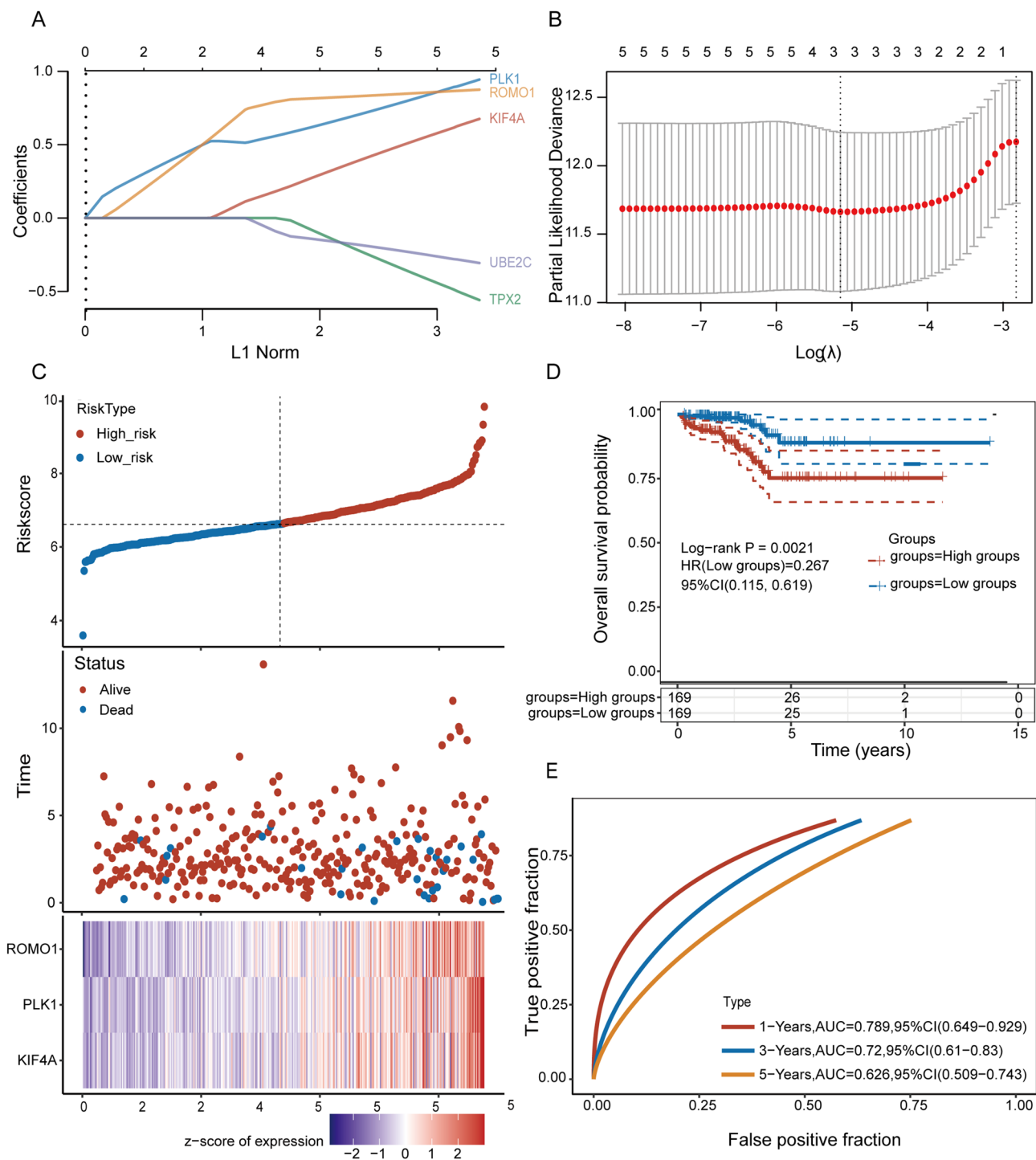


Figure 9. Construction of risk signature prognostic classifier in the TCGA cohort. (A, B). Determining the number of key factors through LASSO analysis. (C). Distribution of Pca patients based on the risk score in the TCGA database. PCA plot for OCs based on the risk score. The survival status for each patient (low-risk population: on the left side; high-risk population: on the right side). (D). Kaplan–Meier curves for the OS of patients in the high- and low-risk groups. (E). ROC curves of hub genes for predicting 1/3/5-year survival in the TCGA dataset.

Kinesin superfamily protein 4A (KIF4A) is found in all eukaryotes and belongs to a family of KIFs that are highly conserved⁷⁶. KIF4A has important roles in DNA repair, DNA replication, spindle organization, cytoplasmic division, and intracellular transport⁷⁷. KIF4A has been previously reported to be aberrantly expressed in many cancers, revealing its function and role in different tumors^{78–80}. KIF4A can promote PCa cell growth through AR and AR-V7-dependent signaling⁷⁸. In the present study, KIF4A expression in different subtypes of

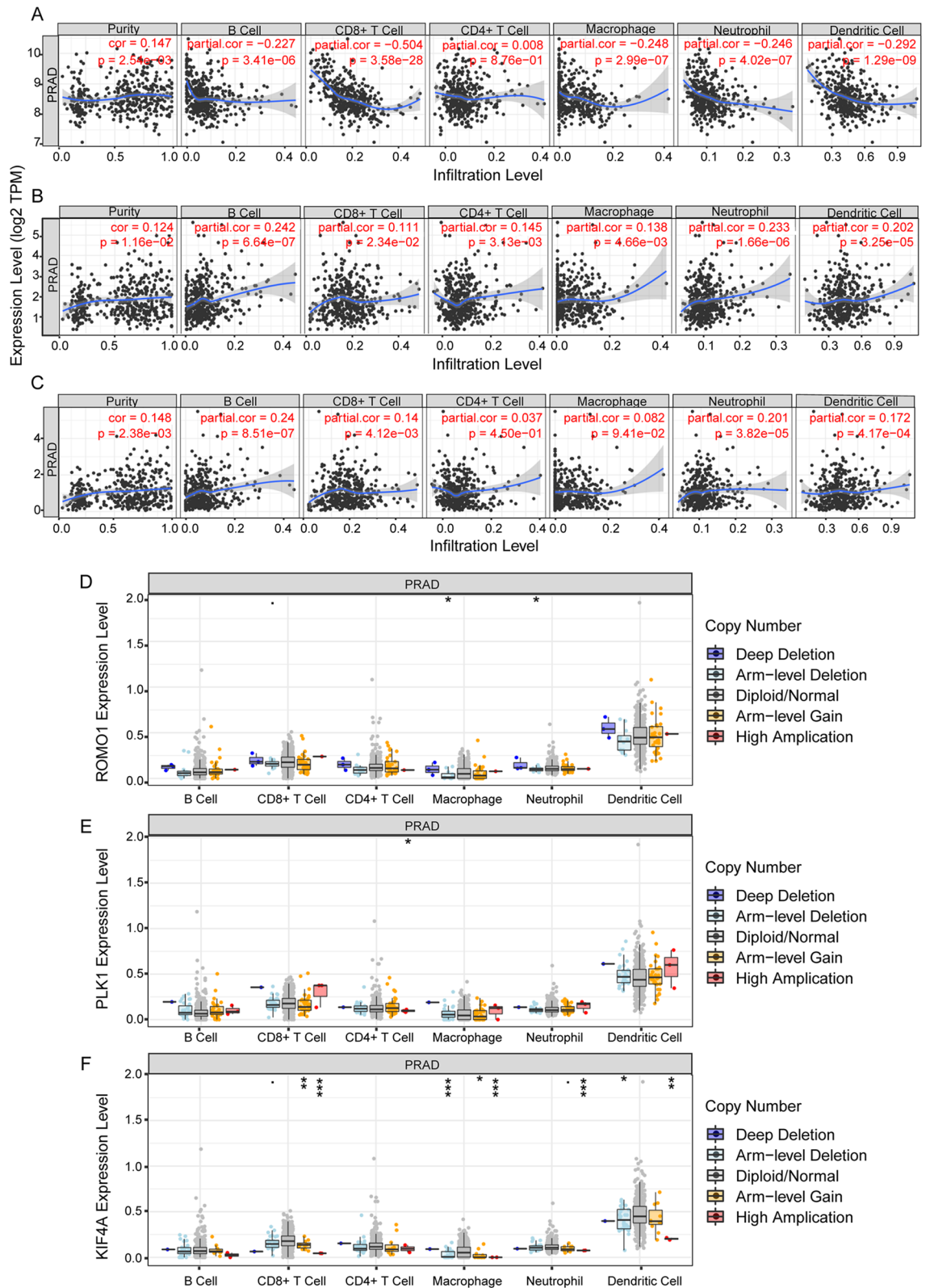


Figure 10. Immune relevance of three hub genes in PCA based on Tumor Immune Estimation Resource (TIMER). (A–C) Correlation study of ROMO1, PLK1 and KIF4A expression with six immune infiltrations, including CD4+ T cells, CD8+ T cells, B cells, macrophages, Neutrophils, and dendritic cells. (D–F) Boxplot shows the comparison of the tumor infiltration levels of ROMO1, PLK1 and KIF4A in immune cells with copy number variants in four categories. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

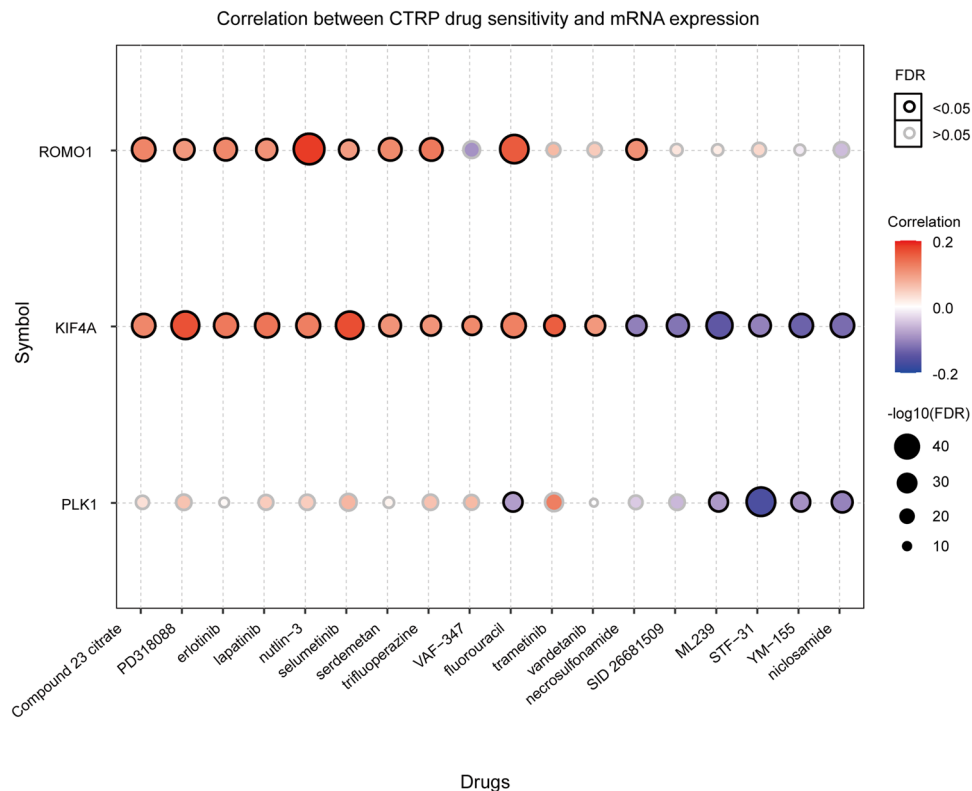


Figure 11. Correlation between CTRP drug sensitivity and mRNA expression.

the prostate was also significantly different from normal samples, and KIF4A expression was also found to be associated with tumor-infiltrating immune cells.

Studies have shown that ROMO1 is overexpressed in hepatocellular carcinoma, colorectal cancer, and glioma^{81–83} but has not been reported in PCa. In the present study, we selected the Wilcoxon test and found that ROMO1 was highly expressed in tumor tissue and significantly different from normal tissue; we also found that the four identified tumor subtypes were significantly different. The expression of ROMO1 was also found to be associated with tumor-infiltrating immune cells, leading to changes in the tumor microenvironment and further increasing tumor heterogeneity; drug sensitivity analysis revealed that nutlin-3 and fluorouracil could be used as potential therapeutic agents for ROMO1.

In general, tumor pathogenesis involves many interacting signaling pathways, including tumor cell proliferation, cell immortalization, invasion, and migration, etc.⁸⁴. The complexity of cancer can be reflected through the tumor microenvironment, while protein interactions can further increase heterogeneity between tumors^{85,86}. In the present study, we have used weighted gene co-expression network analysis (WGCNA), which classifies the gene co-expression network of PCa into ten highly correlated signature modules. The modules were then correlated with specific clinical features to identify genes that are key to tumorigenesis and transformation, to help identify potential mechanisms involved, and to explore candidate biomarkers. However, WGCNA has the limitation of not being able to coexist at different levels of clustering within a single network, thus not reflecting the multi-scale hierarchical nature of complex networks. Multi-scale embedded gene co-expression network analysis (MEGENA), on the other hand, allows the construction and analysis of large-scale planar filtered co-expression networks to the greatest extent possible⁵³. Parallelization of embedded network construction and the identification of multiscale clustering structures are two key components of MEGENA, which is an essential complement to existing co-expression network analysis methods by identifying multiscale modular systems and co-expression networks with varying degrees of sparse and tight connectivity. Here, by using WGCNA in conjunction with MEGENA to construct a gene co-expression network for PCa, we identified more meaningful clusters of co-expressed genes and identified key biomarkers associated with prostate carcinogenesis in transformation.

In this study, we combined various bioinformatic analysis methods, especially the introduction of weighted gene co-expression network analysis and multi-scale chimeric network analysis, to reveal that ROMO1 may serve as a new key prognostic marker for PCa. However, the article still has some limitations. Firstly, the role of ROMO1 has not been validated experimentally in vivo and in vitro, and secondly, the fact that the number of cancer samples and normal samples are not identical to each other has led to some preference in our data. We believe that if we had direct access to a larger sample of clinical sequencing data and sample information, we would obtain better and more accurate results.

Conclusions

In conclusion, our systematic analysis of the TCGA database supported by array-based and sequence-based PCA data has identified ROMO1, a key gene closely associated with the PCA tumor microenvironment, and the essential signaling pathways involved. For this reason, we propose that ROMO1 may serve as a potential biomarker and therapeutic target for PCA. It may provide new theoretical inferences for the diagnosis and prognosis of PCA.

Data availability

All data are available. Please contact us to access if it is needed.

Received: 4 September 2021; Accepted: 7 December 2021

Published online: 07 January 2022

References

- Ge, R. *et al.* Epigenetic modulations and lineage plasticity in advanced prostate cancer. *Ann. Oncol.* **31**, 470–479. <https://doi.org/10.1016/j.annonc.2020.02.002> (2020).
- Auchus, R. J. & Sharifi, N. Sex hormones and prostate cancer. *Ann. Rev. Med.* **71**, 33–45. <https://doi.org/10.1146/annurev-med-051418-060357> (2020).
- Chen, R. C. *et al.* Association between choice of radical prostatectomy, external beam radiotherapy, brachytherapy, or active surveillance and patient-reported quality of life among men with localized prostate cancer. *JAMA* **317**, 1141–1150. <https://doi.org/10.1001/jama.2017.1652> (2017).
- Siegel, D. A., O'Neil, M. E., Richards, T. B., Dowling, N. F. & Weir, H. K. Prostate cancer incidence and survival, by stage and race/ethnicity - United States, 2001–2017. *MMWR Morb. Mortal Wkly Rep.* **69**, 1473–1480. <https://doi.org/10.15585/mmwr.mm6941a1> (2020).
- Bourke, L. *et al.* Exercise for men with prostate cancer: A systematic review and meta-analysis. *Eur Urol* **69**, 693–703. <https://doi.org/10.1016/j.eururo.2015.10.047> (2016).
- Venkateswaran, V. & Klotz, L. H. Diet and prostate cancer: Mechanisms of action and implications for chemoprevention. *Nat. Rev. Urol.* **7**, 442–453. <https://doi.org/10.1038/nrurol.2010.102> (2010).
- Wallis, C. J. D. *et al.* Surgery versus radiotherapy for clinically-localized prostate cancer: A systematic review and meta-analysis. *Eur. Urol.* **70**, 21–30. <https://doi.org/10.1016/j.eururo.2015.11.010> (2016).
- Costello, A. J. Considering the role of radical prostatectomy in 21st century prostate cancer care. *Nat. Rev. Urol.* **17**, 177–188. <https://doi.org/10.1038/s41585-020-0287-y> (2020).
- Zaorsky, N. G. *et al.* The evolution of brachytherapy for prostate cancer. *Nat. Rev. Urol.* **14**, 415–439. <https://doi.org/10.1038/nrurol.2017.76> (2017).
- Mottet, N. *et al.* EAU-ESTRO-SIOG guidelines on prostate cancer: Part 1—Screening, diagnosis, and local treatment with curative intent. *Eur. Urol.* **71**, 618–629. <https://doi.org/10.1016/j.eururo.2016.08.003> (2017).
- Widmark, A. *et al.* Endocrine treatment, with or without radiotherapy, in locally advanced prostate cancer (SPCG-7/SFUO-3): an open randomised phase III trial. *Lancet* **373**, 301–308. [https://doi.org/10.1016/S0140-6736\(08\)61815-2](https://doi.org/10.1016/S0140-6736(08)61815-2) (2009).
- Aparicio, A. M. *et al.* Platinum-based chemotherapy for variant castrate-resistant prostate cancer. *Clin. Cancer Res.* **19**, 3621–3630. <https://doi.org/10.1158/1078-0432.CCR-12-3791> (2013).
- Balk, S. P., Ko, Y. J. & Bubley, G. J. Biology of prostate-specific antigen. *J. Clin. Oncol.* **21**, 383–391. <https://doi.org/10.1200/JCO.2003.02.083> (2003).
- Muenst, S. *et al.* The immune system and cancer evasion strategies: therapeutic concepts. *J. Intern. Med.* **279**, 541–562. <https://doi.org/10.1111/joim.12470> (2016).
- Perez, C. R. & De Palma, M. Engineering dendritic cell vaccines to improve cancer immunotherapy. *Nat. Commun.* **10**, 5408. <https://doi.org/10.1038/s41467-019-13368-y> (2019).
- McLaughlin, J. *et al.* Quantitative assessment of the heterogeneity of PD-L1 expression in non-small-cell lung cancer. *JAMA Oncol.* **2**, 46–54. <https://doi.org/10.1001/jamaoncol.2015.3638> (2016).
- Ebert, L. M., Yu, W., Gargett, T. & Brown, M. P. Logic-gated approaches to extend the utility of chimeric antigen receptor T-cell technology. *Biochem. Soc. Trans.* **46**, 391–401. <https://doi.org/10.1042/BST20170178> (2018).
- Bilusic, M., Madan, R. A. & Gulley, J. L. Immunotherapy of prostate cancer: Facts and hopes. *Clin. Cancer Res.* **23**, 6764–6770. <https://doi.org/10.1158/1078-0432.CCR-17-0019> (2017).
- Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94. <https://doi.org/10.1038/nrclinonc.2017.166> (2018).
- Cacchione, S., Biroccio, A. & Rizzo, A. Emerging roles of telomeric chromatin alterations in cancer. *J. Exp. Clin. Cancer Res.* **38**, 21. <https://doi.org/10.1186/s13046-019-1030-5> (2019).
- Kuppers, R. & Stevenson, F. K. Critical influences on the pathogenesis of follicular lymphoma. *Blood* **131**, 2297–2306. <https://doi.org/10.1182/blood-2017-11-764365> (2018).
- Martincorena, I. & Campbell, P. J. Somatic mutation in cancer and normal cells. *Science* **349**, 1483–1489. <https://doi.org/10.1126/science.aab4082> (2015).
- Sahin, U. & Tureci, O. Personalized vaccines for cancer immunotherapy. *Science* **359**, 1355–1360. <https://doi.org/10.1126/science.aar7112> (2018).
- Lee, E., Yong, R. L., Paddison, P. & Zhu, J. Comparison of glioblastoma (GBM) molecular classification methods. *Semin. Cancer Biol.* **53**, 201–211. <https://doi.org/10.1016/j.semcancer.2018.07.006> (2018).
- Wang, Q. *et al.* Tumor evolution of glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment. *Cancer Cell* **32**, 42–56. <https://doi.org/10.1016/j.ccell.2017.06.003> (2017).
- Jackstadt, R. *et al.* Epithelial NOTCH signaling rewires the tumor microenvironment of colorectal cancer to drive poor-prognosis subtypes and metastasis. *Cancer Cell* **36**, 319–333. <https://doi.org/10.1016/j.ccell.2019.08.003> (2019).
- Ulz, P. *et al.* Inference of transcription factor binding from cell-free DNA enables tumor subtype prediction and early detection. *Nat Commun* **10**, 4666. <https://doi.org/10.1038/s41467-019-12714-4> (2019).
- Kris, M. G. *et al.* Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. *JAMA* **311**, 1998–2006. <https://doi.org/10.1001/jama.2014.3741> (2014).
- Chow, R. D. *et al.* AAV-mediated direct in vivo CRISPR screen identifies functional suppressors in glioblastoma. *Nat Neurosci* **20**, 1329–1341. <https://doi.org/10.1038/nn.4620> (2017).
- Cancer Genome Atlas Research N. The molecular taxonomy of primary prostate cancer. *Cell* **163**, 1011–1025. <https://doi.org/10.1016/j.cell.2015.10.025> (2015).
- Wang, P. *et al.* Macrophage achieves self-protection against oxidative stress-induced ageing through the Mst-Nrf2 axis. *Nat. Commun.* **10**, 755. <https://doi.org/10.1038/s41467-019-08680-6> (2019).

32. Lee, S., Park, Y. H., Chung, J. S. & Yoo, Y. D. Romo1 and the NF-kappaB pathway are involved in oxidative stress-induced tumor cell invasion. *Int. J. Oncol.* **46**, 2021–2028. <https://doi.org/10.3892/ijo.2015.2889> (2015).
33. Srinivas, U. S., Tan, B. W. Q., Vellayappan, B. A. & Jeyasekharan, A. D. ROS and the DNA damage response in cancer. *Redox. Biol.* **25**, 101084. <https://doi.org/10.1016/j.redox.2018.101084> (2019).
34. Idelchik, M., Begley, U., Begley, T. J. & Melendez, J. A. Mitochondrial ROS control of cancer. *Semin. Cancer Biol.* **47**, 57–66. <https://doi.org/10.1016/j.semcancer.2017.04.005> (2017).
35. Jin, M. *et al.* MCUR1 facilitates epithelial-mesenchymal transition and metastasis via the mitochondrial calcium dependent ROS/Nrf2/Notch pathway in hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* **38**, 136. <https://doi.org/10.1186/s13046-019-1135-x> (2019).
36. Sosa, V. *et al.* Oxidative stress and cancer: An overview. *Ageing Res. Rev.* **12**, 376–390. <https://doi.org/10.1016/j.arr.2012.10.004> (2013).
37. Lee, S. B. *et al.* Romo1 is a negative-feedback regulator of Myc. *J. Cell. Sci.* **124**, 1911–1924. <https://doi.org/10.1242/jcs.079996> (2011).
38. Liu, Y. *et al.* High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell* **183**, 1665–1681e1618. <https://doi.org/10.1016/j.cell.2020.10.026> (2020).
39. Grindberg, R. V. *et al.* RNA-sequencing from single nuclei. *Proc. Natl. Acad. Sci. U S A* **110**, 19802–19807. <https://doi.org/10.1073/pnas.1319700110> (2013).
40. Hutter, C. & Zenklusen, J. C. The cancer genome atlas: Creating lasting value beyond its data. *Cell* **173**, 283–285. <https://doi.org/10.1016/j.cell.2018.03.042> (2018).
41. Tang, Z. *et al.* GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **45**, W98–w102. <https://doi.org/10.1093/nar/gkx247> (2017).
42. Thul, P. J. & Lindskog, C. The human protein atlas: A spatial map of the human proteome. *Protein Sci.* **27**, 233–244. <https://doi.org/10.1002/pro.3307> (2018).
43. Goldman, M. J. *et al.* Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* **38**, 675–678. <https://doi.org/10.1038/s41587-020-0546-8> (2020).
44. Lancichinetti, A. & Fortunato, S. Consensus clustering in complex networks. *Sci. Rep.* **2**, 336. <https://doi.org/10.1038/srep00336> (2012).
45. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573. <https://doi.org/10.1093/bioinformatics/btq170> (2010).
46. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47. <https://doi.org/10.1093/nar/gkv007> (2015).
47. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523. <https://doi.org/10.1038/s41467-019-09234-6> (2019).
48. van Dam, S., Vosa, U., van der Graaf, A., Franke, L. & de Magalhaes, J. P. Gene co-expression analysis for functional classification and gene-disease predictions. *Brief Bioinform.* **19**, 575–592. <https://doi.org/10.1093/bib/bbw139> (2018).
49. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489. <https://doi.org/10.1038/nature10523> (2011).
50. Thorsson, V. *et al.* The immune landscape of cancer. *Immunity* **48**, 812–830e814. <https://doi.org/10.1016/j.immuni.2018.03.023> (2018).
51. Prasetyanti, P. R. & Medema, J. P. Intra-tumor heterogeneity from a cancer stem cell perspective. *Mol Cancer* **16**, 41. <https://doi.org/10.1186/s12943-017-0600-4> (2017).
52. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559. <https://doi.org/10.1186/1471-2105-9-559> (2008).
53. Song, W. M. & Zhang, B. Multiscale embedded gene co-expression network analysis. *PLoS Comput Biol* **11**, e1004574. <https://doi.org/10.1371/journal.pcbi.1004574> (2015).
54. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361. <https://doi.org/10.1093/nar/gkw1092> (2017).
55. da Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57. <https://doi.org/10.1038/nprot.2008.211> (2009).
56. Szklarczyk, D. *et al.* STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613. <https://doi.org/10.1093/nar/gky1131> (2019).
57. Chin, C. H. *et al.* cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **8**(Suppl4), S11. <https://doi.org/10.1186/1752-0509-8-S4-S11> (2014).
58. Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **47**, W556–W560. <https://doi.org/10.1093/nar/gkz430> (2019).
59. Colwill, K. & Graslund, S. A roadmap to generate renewable protein binders to the human proteome. *Nat. Methods* **8**, 551–558. <https://doi.org/10.1038/nmeth.1607> (2011).
60. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* **33**, 1–22 (2010).
61. Li, T. *et al.* TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res* **77**, e108–e110. <https://doi.org/10.1158/0008-5472.CAN-17-0307> (2017).
62. Liu, C. J. *et al.* GSCALite: a web server for gene set cancer analysis. *Bioinformatics* **34**, 3771–3772. <https://doi.org/10.1093/bioinformatics/bty411> (2018).
63. Dean, C. B. & Nielsen, J. D. Generalized linear mixed models: a review and some extensions. *Lifetime Data Anal* **13**, 497–512. <https://doi.org/10.1007/s10985-007-9065-x> (2007).
64. Gorchakov, A. A., Kulemzin, S. V., Kochneva, G. V. & Tarantin, A. V. Challenges and prospects of chimeric antigen receptor T-cell therapy for metastatic prostate cancer. *Eur Urol* **77**, 299–308. <https://doi.org/10.1016/j.eururo.2019.08.014> (2020).
65. Trock, B. J. *et al.* Prostate cancer-specific survival following salvage radiotherapy vs observation in men with biochemical recurrence after radical prostatectomy. *JAMA* **299**, 2760–2769. <https://doi.org/10.1001/jama.299.23.2760> (2008).
66. Mateo, J. *et al.* Managing nonmetastatic castration-resistant prostate cancer. *Eur Urol* **75**, 285–293. <https://doi.org/10.1016/j.eururo.2018.07.035> (2019).
67. Chang, A. J., Autio, K. A., Roach, M. 3rd. & Scher, H. I. High-risk prostate cancer-classification and therapy. *Nat Rev Clin Oncol* **11**, 308–323. <https://doi.org/10.1038/nrclinonc.2014.68> (2014).
68. Wang, G., Zhao, D., Spring, D. J. & DePinho, R. A. Genetics and biology of prostate cancer. *Genes Dev* **32**, 1105–1140. <https://doi.org/10.1101/gad.315739.118> (2018).
69. Zhao, S. G. *et al.* The Immune Landscape of Prostate Cancer and Nomination of PD-L2 as a Potential Therapeutic Target. *J Natl Cancer Inst* **111**, 301–310. <https://doi.org/10.1093/jnci/djy141> (2019).
70. Steegmaier, M. *et al.* BI 2536, a potent and selective inhibitor of polo-like kinase 1, inhibits tumor growth in vivo. *Curr Biol* **17**, 316–322. <https://doi.org/10.1016/j.cub.2006.12.037> (2007).
71. Frank, T. *et al.* Cell cycle arrest in mitosis promotes interferon-induced necroptosis. *Cell Death Differ* **26**, 2046–2060. <https://doi.org/10.1038/s41418-019-0298-5> (2019).

72. Smits, V. A. *et al.* Polo-like kinase-1 is a target of the DNA damage checkpoint. *Nat Cell Biol* **2**, 672–676. <https://doi.org/10.1038/35023629> (2000).
73. Macurek, L. *et al.* Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery. *Nature* **455**, 119–123. <https://doi.org/10.1038/nature07185> (2008).
74. Nieto-Jimenez, C. *et al.* Inhibition of the mitotic kinase PLK1 overcomes therapeutic resistance to BET inhibitors in triple negative breast cancer. *Cancer Lett* **491**, 50–59. <https://doi.org/10.1016/j.canlet.2020.06.020> (2020).
75. Zhang, Z. *et al.* Plk1 inhibition enhances the efficacy of androgen signaling blockade in castration-resistant prostate cancer. *Cancer Res* **74**, 6635–6647. <https://doi.org/10.1158/0008-5472.CAN-14-1916> (2014).
76. Miki, H., Okada, Y. & Hirokawa, N. Analysis of the kinesin superfamily: Insights into structure and function. *Trends Cell Biol* **15**, 467–476. <https://doi.org/10.1016/j.tcb.2005.07.006> (2005).
77. Sheng, L., Hao, S. L., Yang, W. X. & Sun, Y. The multiple functions of kinesin-4 family motor protein KIF4 and its clinical potential. *Gene* **678**, 90–99. <https://doi.org/10.1016/j.gene.2018.08.005> (2018).
78. Cao, Q. *et al.* Targeting the KIF4A/AR axis to reverse endocrine therapy resistance in castration-resistant prostate cancer. *Clin Cancer Res* **26**, 1516–1528. <https://doi.org/10.1158/1078-0432.CCR-19-0396> (2020).
79. Hu, G. *et al.* FOXM1 promotes hepatocellular carcinoma progression by regulating KIF4A expression. *J Exp Clin Cancer Res* **38**, 188. <https://doi.org/10.1186/s13046-019-1202-3> (2019).
80. Taniwaki, M. *et al.* Activation of KIF4A as a prognostic biomarker and therapeutic target for lung cancer. *Clin Cancer Res* **13**, 6624–6631. <https://doi.org/10.1158/1078-0432.CCR-07-1328> (2007).
81. Chung, J. S. *et al.* Overexpression of Romo1 promotes production of reactive oxygen species and invasiveness of hepatic tumor cells. *Gastroenterology* **143**, 1084–1094. <https://doi.org/10.1053/j.gastro.2012.06.038> (2012).
82. Kim, H. J. *et al.* Reactive oxygen species modulator-1 (Romo1) predicts unfavorable prognosis in colorectal cancer patients. *PLoS ONE* **12**, e0176834. <https://doi.org/10.1371/journal.pone.0176834> (2017).
83. Yu, M. O. *et al.* Romo1 is associated with ROS production and cellular growth in human gliomas. *J Neurooncol* **121**, 73–81. <https://doi.org/10.1007/s11060-014-1608-x> (2015).
84. Rodriguez, F. J., Folpe, A. L., Giannini, C. & Perry, A. Pathology of peripheral nerve sheath tumors: Diagnostic overview and update on selected diagnostic problems. *Acta Neuropathol* **123**, 295–319. <https://doi.org/10.1007/s00401-012-0954-z> (2012).
85. Boedtker, E. & Pedersen, S. F. The acidic tumor microenvironment as a driver of cancer. *Annu Rev Physiol* **82**, 103–126. <https://doi.org/10.1146/annurev-physiol-021119-034627> (2020).
86. Martinez-Outschoorn, U. E., Peiris-Pages, M., Pestell, R. G., Sotgia, F. & Lisanti, M. P. Cancer metabolism: A therapeutic perspective. *Nat Rev Clin Oncol* **14**, 11–31. <https://doi.org/10.1038/nrclinonc.2016.60> (2017).

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (U1804179) and the Nanhu Scholars Program of XYNU. This work was also supported by the Graduate Research Innovation Foundation of Xinyang Normal University (Nos. 2020KYJJ36, 2020KYJJ60).

Author contributions

L.W., W.L., and X.S. conceived and designed the study. M.F., J.Y. and Y.M. carried out experiments and statistical analysis. Y.Z., Y.M., J.L., R.L., X.Z., P.Z., L.Z., J.F., and Y.C., provided some critical revision; X.L., Y.W. and Z.L. wrote the manuscript. The author(s) read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-03946-w>.

Correspondence and requests for materials should be addressed to L.W., W.L. or X.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022