

ORIGINAL ARTICLE

Identification of castration-resistant prostate cancer-related hub genes using weighted gene co-expression network analysis

Yifei Cheng^{1,2} | Lu Li³ | Zongshi Qin⁴ | Xiao Li¹ | Feng Qi^{1,2} 

¹Department of Urologic Surgery, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Affiliated Cancer Hospital of Nanjing Medical University, Nanjing, China

²Department of Urology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

³Nanjing Medical University, Nanjing, China

⁴School of Chinese Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China

Correspondence

Xiao Li, Department of Urologic Surgery, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Affiliated Cancer Hospital of Nanjing Medical University. Email: leex91@163.com

Feng Qi, Department of Urologic Surgery, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Affiliated Cancer Hospital of Nanjing Medical University, Nanjing 210009, China; Department of Urology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210009, China. Email: qf199408@163.com

Funding information

National Natural Science Foundation of China, Grant/Award Number: 81702520; Research Project of Jiangsu Cancer Hospital, Grant/Award Number: ZN201602; Medical Research Project of Jiangsu Provincial Health and Family Planning Commission, Grant/Award Number: H2018052; Young Talents Program of Jiangsu Cancer Hospital, Grant/Award Number: 2017YQL-04

Abstract

Prostate cancer is the most common malignancy in urinary system and brings heavy burdens in men. We downloaded gene expression profile of mRNA and related clinical data of GSE70768 data set from public database. Weighted gene co-expression network analysis (WGCNA) was used to identify the relationships between gene modules and clinical features, as well as the candidate genes. Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analyses were developed to investigate the potential functions of related hub genes. Importantly, basic experiments were performed to verify the relationship between hub genes and the phenotype previously identified. Lastly, copy number variation (CNV) analysis was conducted to explore the genetical alteration. WGCNA identified that black module was the most relevant module which was tightly related to castration-resistant prostate cancer (CRPC) phenotype. KEGG and GO analysis results revealed genes in black module were mainly related to RNA splicing. Additionally, 9 genes were chosen as hub genes and heterogeneous nuclear ribonucleoprotein A2/B1 (*HNRNPA2B1*), golgin A8 family member B (*GOLGA8B*) and mitogen-activated protein kinase 8 interacting protein 3 (*MAPK8IP3*) were identified to be associated with PCa progression and prognosis. Moreover, all above three genes were highly expressed in CRPC-like cells and their suppression led to hindered cell proliferation in vitro. Finally, CNV analysis found that amplification was the main type of alteration of the 3 hub genes. Our study found that *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* were identified to be tightly associated with tumour progression and prognosis, and further researches are needed before clinical application.

KEYWORDS

CRPC, hub genes, prognosis, WGCNA

Yifei Cheng, Lu Li and Zongshi Qin contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Journal of Cellular and Molecular Medicine* published by Foundation for Cellular and Molecular Medicine and John Wiley & Sons Ltd

1 | INTRODUCTION

Prostate cancer (PCa) is the most common malignancy in urinary system, leading to the fifth cause of death in male worldwide. In 2018, the global estimated new cases and deaths were 1 276 106 and 358 989, respectively.¹ Localized PCa patients can be treated effectively via various methods including active surveillance, watchful waiting, surgery and radiation. However, management for advanced diseases remains to be a difficult problem for urologists. Generally, androgen deprivation therapy (ADT), raised by Charles Huggins² firstly in the 1940s, is the standard treatment for advanced PCa patients. Types of ADT forms mainly include medical castration (luteinizing hormone-releasing hormone (LHRH) agonist) and surgical castration (bilateral orchiectomy). Unfortunately, ADT therapy for these castration-resistant prostate cancer (CRPC) patients is not curative. As a result, progressing to CRPC condition is inevitable within two years.³⁻⁵

Taken together, previous data indicated that 10%-20% of all PCa patients develop CRPC within approximately 5 years from initial diagnosis. More than 84% of CRPC patients had metastases at diagnosis with bone as the most preferential site.⁶ For those who had no metastases at diagnosis of CRPC, 33% of patients were estimated to develop metastases within 2 years.⁷ Radionuclide therapy, radiotherapy, bisphosphonates and opioids were the most common treatment types for those patients with metastatic CRPC and skeletal symptoms. A systematic review showed that only about one-third of CRPC patients received chemotherapy while the rest could only receive supportive care or steroids.⁸ Prognosis of CRPC patients is poor, especially for those patients with metastatic CRPC. Study conducted by Emmanuel S. Antonarakis et al showed that the median survival from CRPC initial diagnosis was 14 months while the overall survival of patients with metastatic CRPC was only 9-13 months.^{8,9}

Nowadays, the specific mechanisms underlying the development of castration resistance have not been well discussed and strategies of CRPC treatment are limited. Study conducted by Wu et al¹⁰ showed that PCa patients with variant allele of *HSD3B1* increased the risk of progressing to CRPC, but it was not associated with biochemical recurrence, mortality or overall survival. Joshua W. Russo et al¹¹ proposed that down-regulation of dipeptidyl peptidase 4 promoted progression to CRPC. With the development of DNA microarray and high-throughput sequencing, it is convenient to explore hub genes related to tumour process or progression benefited from bioinformatics techniques and big data integration. Weighted gene co-expression network analysis (WGCNA) is a powerful biological method which can systematically investigate highly synergistically altered gene modules. It can thoroughly study the relationship between gene sets and clinical characteristics, and identify possible biomarkers for further researches. In this study, we aimed to find and validate differentially expressed genes (DEGs) which were associated with castration resistance in PCa by WGCNA. In addition, Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analyses were developed to investigate the potential functions of hub genes in the key module.

2 | METHODS AND MATERIALS

2.1 | Data processing

A flow chart of this study was shown in Figure 1. Gene expression profile of mRNA and related clinical data of PCa patients were downloaded from Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>). GSE70768 included 16 331 genes and 125 tumour tissues, which was used to construct WGCNA for this

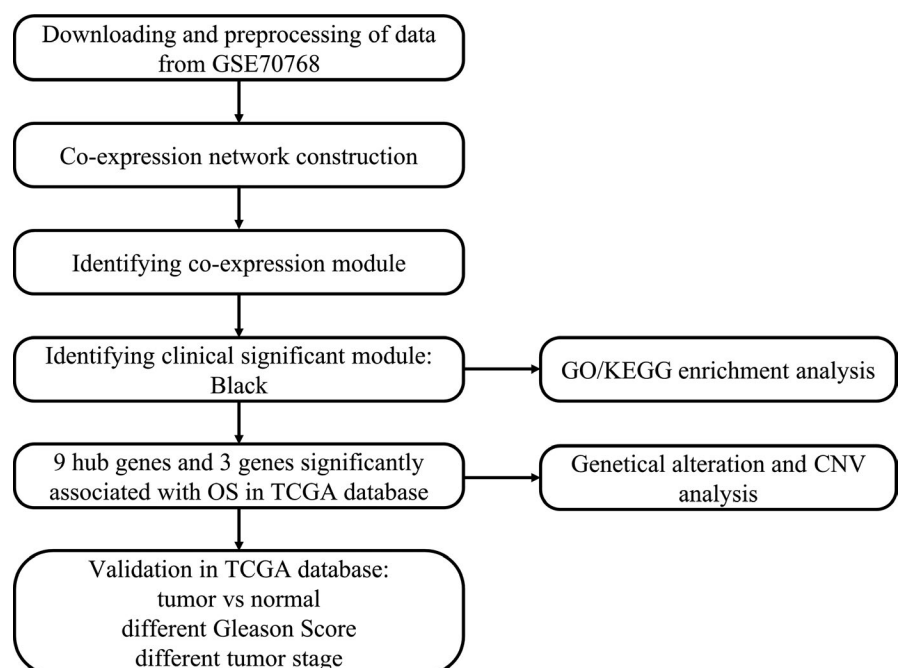


FIGURE 1 Flow chart of data preparation, processing, analysis and validation

study. Among them, 13 were CRPC tissues and 112 were hormone-sensitive prostate cancer (HSPC) tissues. The Cancer Genome Atlas (TCGA) data portal (<https://portal.gdc.cancer.gov/>) was used for validation of hub genes. The data of mRNA expression were transformed to values in transcripts per million (TPM).

2.2 | Construction of the co-expression network

'WGCNA' R package was applied to find clinical traits-related modules and hub genes.¹² The formulas used to calculate the interaction coefficient between genes and transform the connection coefficient into a weighted coefficient were specifically described in previous studies.¹³ The top 5000 most variant genes by analysis of variance were retained for WGCNA analysis.

Firstly, we use `pickSoftThreshold` function to find a soft threshold power β in accordance with standard scale-free networks. Second, the adjacencies between all filtered genes were calculated using the power adjacent function to Pearson's correlation matrix to transform data into a topological overlap matrix (TOM), and the corresponding dissimilarity (1-TOM) was calculated. Then, the dissimilarity of module eigengenes was calculated to further analyse the module. Finally, we set soft-thresholding power of 9, cut-off height of 0.25 and minimal module size of 30 to identify key modules.

2.3 | Identification of clinically significant modules

The module eigengene (ME) was defined as the first principal component of a given module, which could be regarded as a representative of the gene expression profiles from a module. Gene significance (GS) was defined as the \log_{10} transformation of the P value ($GS = \lg P$) in the linear regression between gene expression and pathological progression. Module significance (MS) represented the average GS for all the genes in a module. In a word, the module with the absolute MS ranked first among all the selected modules was chosen as the one related with clinical trait.

2.4 | Pathway enrichment and GO analysis

DAVID (<http://david.abcc.ncifcrf.gov/>) is a database for annotation, visualization and integrated discovery. KEGG pathway and GO analysis of genes in the module with the highest correlation with clinical traits were carried out using DAVID (version 6.8) online tools: functional annotation. Enriched GO terms and KEGG pathways were identified according to the cut-off criterion of $P < 0.1$.

2.5 | Hub genes detection and validation

Hub genes were defined as those with $\text{cor.geneTraitSignificance} > 0.2$ and $\text{cor.geneModuleMembership} > 0.85$. Then,

TCGA-PRAD data were applied to perform survival analysis to validate and select genes with clinical significance. Then, we compared the expression of genes between tumours and normal tissues along with expression between tumours with different clinical status in TCGA database.

2.6 | Cell culture and transfection

The human androgen-dependent PCa cell line LNCaP (ATCC, Manassas, VA, USA) and three CRPC-like cell lines including LNCaP-AI, LNCaP-Bic and C4-2 (ATCC) were applied in this study. LNCaP and C4-2 cells were maintained in RPMI-1640 medium (Gibco, Waltham, MA, USA) containing 10% foetal bovine serum (FBS; Gibco) and 1% penicillin/streptomycin (Gibco). LNCaP-AI cells were grown in RPMI-1640 without phenol red and supplemented with 10% charcoal-stripped FBS (Gibco), whereas LNCaP-Bic cells were added with 20 μM bicalutamide (Sigma, St. Louis, MO USA). Prior to any experiment, LNCaP cells were maintained in phenol red-free medium containing charcoal-stripped FBS for no less than 48 hours. All cells were incubated under a humidified atmosphere with 5% CO_2 at 37°C.

With respect to cell transfection, LNCaP-AI and C4-2 cells were transfected with indicated short hairpin RNAs (shRNAs), such as sh-HNRNPA2B1#1/2, sh-GOLGA8B#1/2, sh-MAPK8IP3#1/2 and corresponding control sh-NC, by use of Lipofectamine 2000 (Invitrogen, Waltham, MA, USA) in line with manufacturers' guide.

2.7 | Cell proliferation detection

Cell proliferation was analysed via colony formation and EdU (5-ethynyl-2'-deoxyuridine) assays. For Colony formation assay, 1×10^3 cells were seeded into six-well plates with three replicates. Two weeks later, colonies were stained via 0.5% crystal violet and those with more than 50 cells were counted after imaged. For EdU assay, EdU staining kit (Ribobio) was used based on the protocol. In brief, cells in 96-well plates were processed with 50 $\mu\text{mol/L}$ EdU medium (100 μL) for 3 hours. After being fixed and permeabilized, cells were stained via DAPI and then observed and imaged under a fluorescence microscope (Olympus, Tokyo, Japan).

2.8 | Cell apoptosis analysis

Cell apoptosis was examined by TUNEL (terminal deoxynucleotidyl transferase-mediated dUTP nick end labelling) assay and flow cytometry analysis. In regard to TUNEL assay, cells were permeated with 0.1% Triton X-100, rinsed in phosphate buffer saline (PBS) and then treated with TUNEL reaction for 1 hour according to the protocol of the TUNEL staining kit. Finally, apoptotic cells were observed and pictured by Olympus fluorescence microscope. As for flow cytometry analysis, cells were treated with 15 minutes of staining of

FITC-Annexin V and propidium iodide (PI), followed by analysis via flow cytometer (BD Biosciences, Franklin Lakes, NJ, USA).

2.9 | Genetical alteration of hub genes and protein-protein interaction (PPI) analysis

The cBioPortal for Cancer Genomics (<http://www.cbioportal.org/>) is a large-scale cancer genomics database. We select four data sets associated with metastasis (SU2C, MICH, SU2C219 and MPC Project) to explore genetic alterations connected with the 9 hub genes. Next, copy number variation (CNV) data from TCGA database were used to conduct CNV analysis. Samples were classified into three groups: amplification, wild-type and deletion according to the CNV data, and the expression levels of genes were compared among these groups. Finally, we applied String (<https://string-db.org/>) and cystoscope (version 3.7.1) to construct protein-protein interaction network of hub genes.

2.10 | Statistical analyses

Kaplan-Meier curves and the log-rank tests were used to assess association between mRNA expression and overall survival of PCA patients. Relationships between mRNA expression and clinical or CNV features were assessed by the Wilcoxon rank-sum test for comparison between two groups and ANOVA for comparison between more than two groups. All statistical analyses were performed using R (version 3.6.1). All statistical tests were 2-sided, and $P < 0.05$ was considered statistically significant.

3 | RESULTS

3.1 | Weighted co-expression network and Key Modules

Pearson's correlation method and average linkage method were used to cluster the samples from GSE70768 data set (Figure S1). Then, co-expression network was constructed using co-expression analysis. Under the soft-thresholding power of 9, cut height of 0.25 and minimal module size as 30, we identified 13 modules (Figure 2). Later, the heatmaps of the correlations confirmed the independence of these 13 modules and gene expression in each module (Figure 3A-B). After module-trait relationship analysis, we observed that black module had the highest association with castration resistance (Figure 3C). In addition, a strong relationship was detected between genes in black module and CRPC ($P = 6.1 \times 10^{-9}$, Figure 3D). Consequently, we focused on black module and the clinical characteristic of CRPC in our further study.

3.2 | Pathway enrichment and gene ontology analysis

KEGG pathway enrichment and GO analysis were applied to perform gene annotation. Figure 4A exhibited the result of KEGG pathway enrichment. The most significant pathway was spliceosome pathway and the other significant pathways included RNA transport, surveillance and degradation pathways. Biological process (BP) of GO analysis showed genes in the black module were mainly associated with RNA splicing, mRNA processing, cilium morphogenesis and apoptotic process (Figure 4B). In a word, the black module genes were

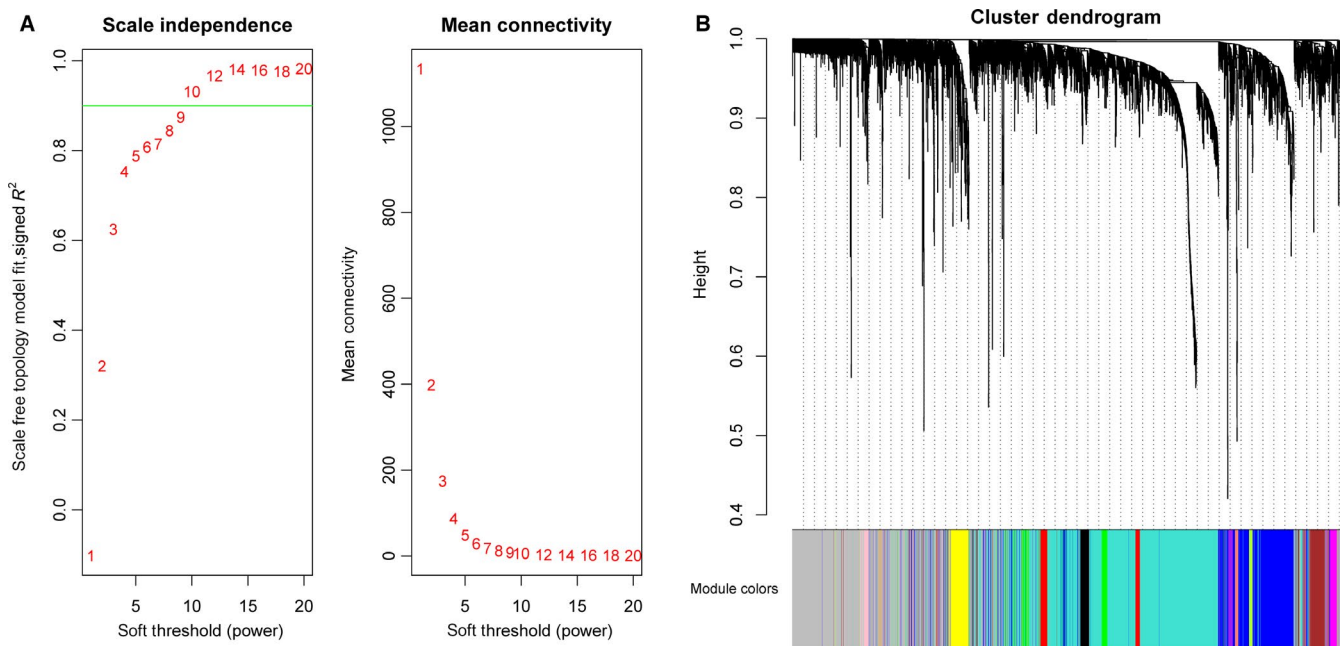


FIGURE 2 Determination of soft-thresholding power in the WGCNA and identification of modules. A, Analysis of the scale-free fit index for various soft-thresholding powers (left). Analysis of the mean connectivity for various soft-thresholding powers (right). B, Dendrogram of all differentially expressed genes was clustered based on a dissimilarity measure (1-TOM). 13 modules were identified

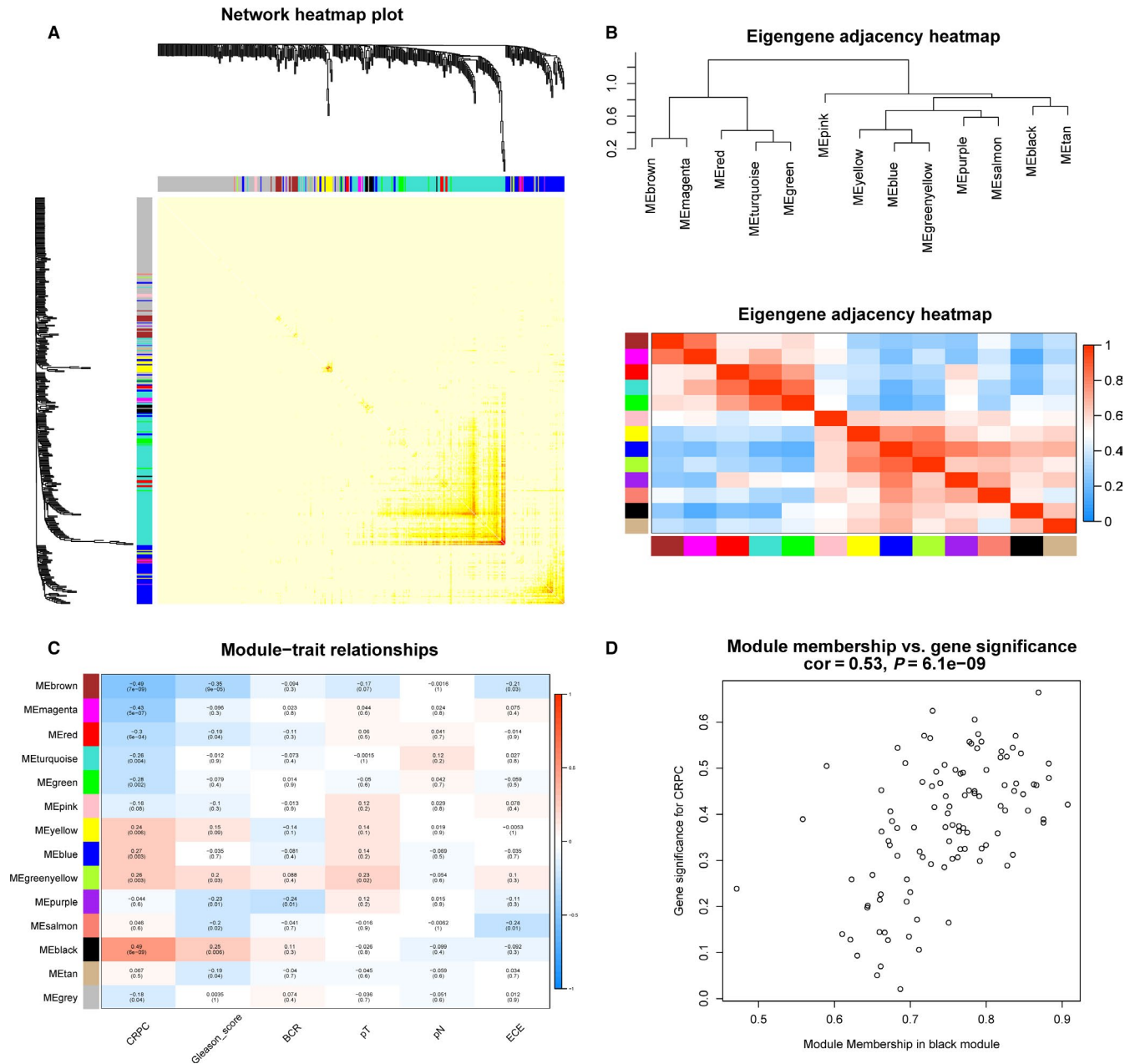


FIGURE 3 Module interaction and identification of modules associated with the clinical traits of PCa. A, The heatmaps of the correlations between 13 modules. B, Clustering of modules eigengenes and heatmaps of the correlations between eigengenes. C, Heatmap of the correlation between module eigengenes and clinical traits of PCa. D, Scatter plot of module eigengenes in the black module

mainly related to the regulation on RNA level involving RNA splicing, processing, transport, surveillance and degradation.

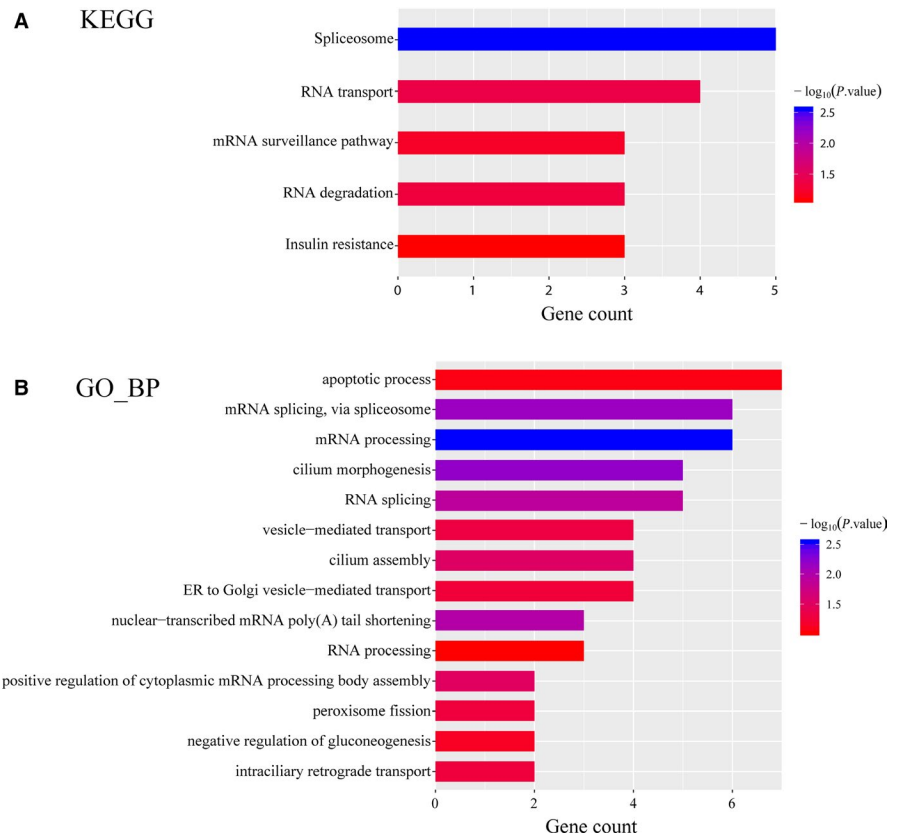
3.3 | Hub genes identification and validation

Nine genes with the high connectivity in black module were chosen as hub genes based on the criteria that $cor.geneModuleMembership > 0.85$ and $cor.geneTraitSignificance > 0.2$. Among them, heterogeneous nuclear ribonucleoprotein A2/B1 (*HNRNPA2B1*), golgin A8 family member B (*GOLGA8B*) and mitogen-activated protein kinase 8 interacting protein 3 (*MAPK8IP3*) were found significantly associated

with overall survival of PCa patients in TCGA database ($P = 4.2 \times 10^{-2}$, 7.3×10^{-4} and 2.2×10^{-2} for *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3*, respectively, Figure 5A-C). The Kaplan-Meier Plotter of the rest 6 genes was shown in Figure S2.

We further evaluated the relationship between these 3 genes and PCa risk. As shown in Figure 5D-F, dramatic differences were found in the expression of these 3 genes between tumours and normal tissues ($P = 3.0 \times 10^{-11}$, 9.4×10^{-7} and 4.1×10^{-4} for *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3*, respectively). In addition, high expression of *HNRNPA2B1* and *MAPK8IP3* was significantly associated with worse pathological N and T stage ($P = 5.3 \times 10^{-5}$ and 3.6×10^{-5} for *HNRNPA2B1* and *MAPK8IP3* in pN1 vs pN0, respectively, $P = 4.1 \times 10^{-5}$ and 2.1×10^{-2} for

FIGURE 4 KEGG pathway and GO analysis of hub genes in black module. A, Bar chart of KEGG pathway analysis for genes in the black module. B, Bar chart of biological process of GO analysis for genes in the black module. KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology



HNRNPA2B1 in pT3 vs pT2 and pT4 vs pT2, respectively. $P = 1.2 \times 10^{-2}$ for *MAPK8IP3* in pT3 vs pT2. Figure 5G-L). Besides, significant differences between high-risk PCa patients [Gleason score (GS) ≥ 8] and low- or medium-risk PCa patients (GS ≤ 6 or =7) were detected in all these 3 genes ($P = 8.5 \times 10^{-6}$ and 5.0×10^{-3} for *HNRNPA2B1* and *MAPK8IP3* in GS ≥ 8 vs GS ≤ 6 , respectively. $P = 2.3 \times 10^{-10}$, 1.8×10^{-4} and 2.2×10^{-10} for *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* in GS ≥ 8 vs GS = 7, respectively. Figure 5M-O). More importantly, we discovered that *HNRNPA2B1* was highly expressed in CRPC-like cell lines and inhibition of *HNRNPA2B1* could hamper CRPC cell proliferation but facilitate apoptosis (Figure 6). As for the expression pattern and function of *GOLGA8B* and *MAPK8IP3* in CRPC-like cells, we observed similar phenomena as that of *HNRNPA2B1* (Figure S3 and S4). In one sentence, *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* might be significant for tumorigenicity of CRPC.

3.4 | Genetical alteration of hub genes and CNV analysis

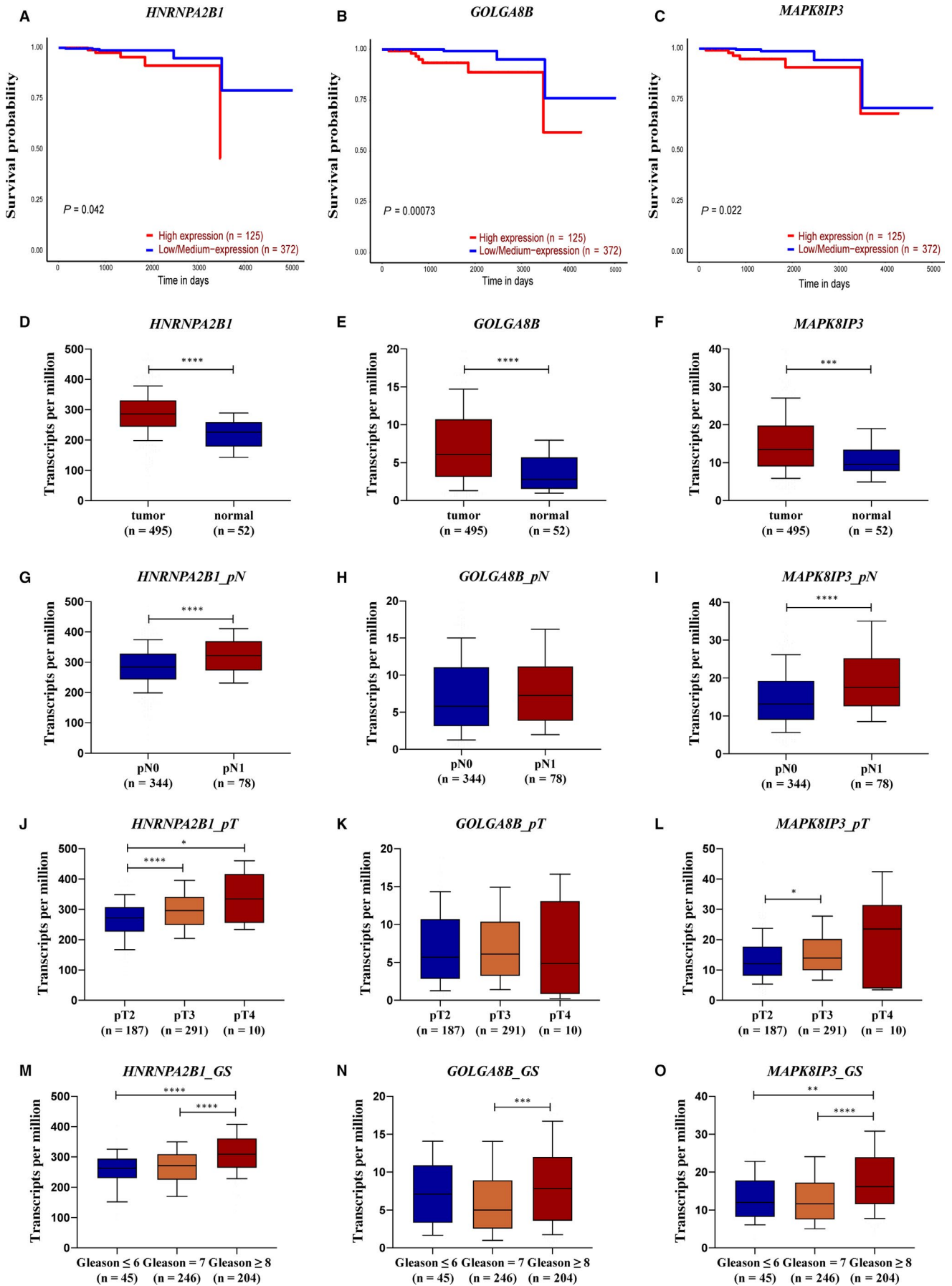
The genetical alteration of the 9 hub genes was analysed using data concerning metastasis from cBioPortal database. The alteration

frequency of each hub gene was shown in Figure 7A. It was shown that *FAM156B*, *HNRNPA2B1* and *MAPK8IP3* altered most (8%, 7% and 7%, respectively, Figure 7A). In addition, we found that amplification was the main type of alteration in all the four subjects (Figure 7B). The CNV analysis of gene *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* confirmed the result that there existed abundant of amplification among this group of genes (Figure 7C-E). Figure S5 showed the relationship of the genes in the black module and the other frequently altered neighbour genes. *HNRNPA2B1* and *RBM25* were significantly important in the network.

4 | DISCUSSION

PCa is a heterogeneous disease, and its pathogenesis remains unclear. Many studies have explored novel biomarkers that related to oncogenesis, tumour progression and prognosis using RNA sequencing and gene chip. However, huge inconsistencies existed between the DEGs found in these studies.¹⁴ Song et al¹⁵ discovered *LMNB1*, *TK1*, *RACGAP1* and *ZWINT* as candidate genes for diagnosis and prognosis of PCa using Robust Rank Aggregation (RRA) method. Hou et al¹⁶ developed a prognostic and diagnostic prediction model

FIGURE 5 Validation of hub genes in black module. A-C, Association between *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* expression and overall survival time in the TCGA-PRAD data set. D-F, Expression difference between tumour and normal tissues of *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* in the TCGA-PRAD data set. G-I, Expression difference between tumour tissues with different pathological N stages of *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* in the TCGA-PRAD data set. J-L, Expression difference between tumour tissues with different pathological T stages of *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* in the TCGA-PRAD data set. (M-O) Expression difference between tumour tissues with Gleason score of ≤ 6 , =7 and ≥ 8 of *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* in the TCGA-PRAD data set



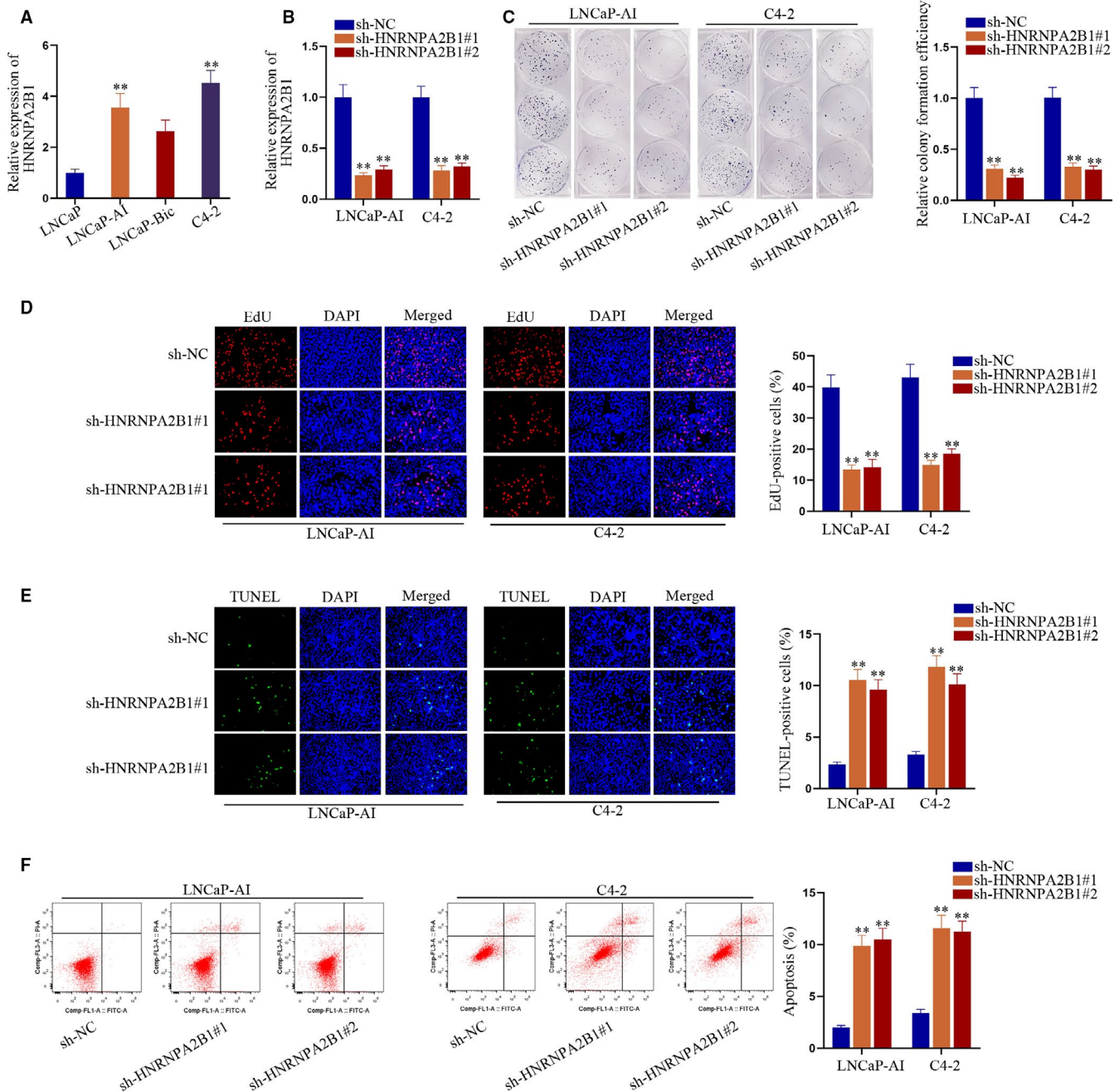


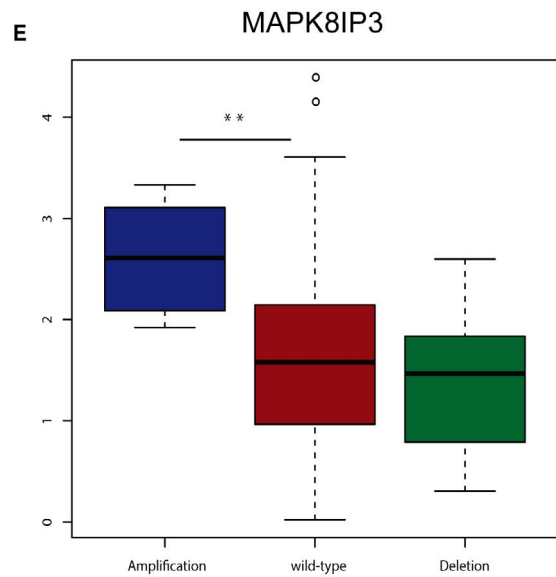
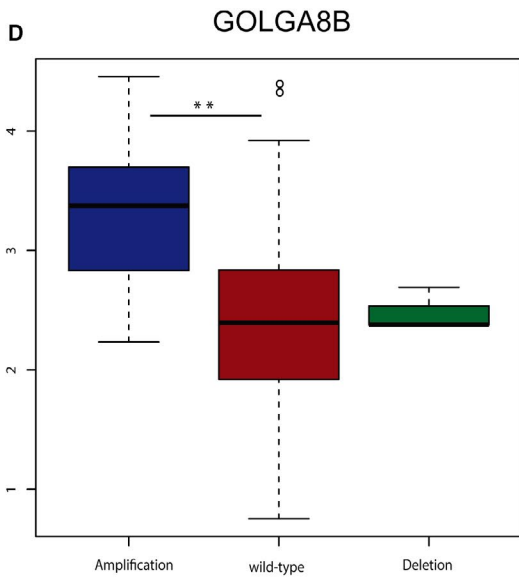
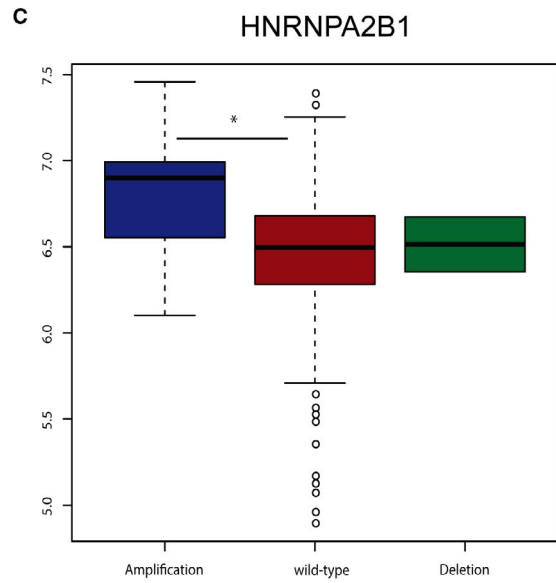
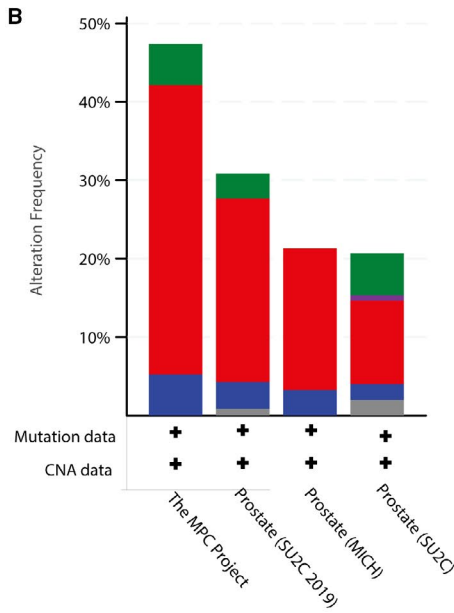
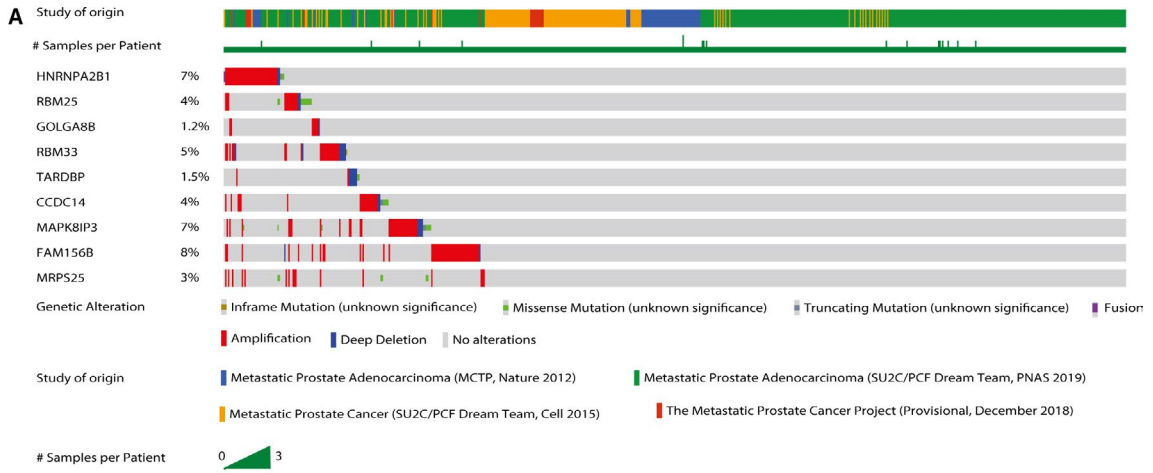
FIGURE 6 Silencing HNRNPA2B1 impaired CRPC cell proliferation and accelerated cell apoptosis in vitro. A, Expression trend of HNRNPA2B1 in CRPC-like cells compared to androgen-dependent LNCaP cells. B, Knockdown efficiencies of sh-HNRNPA2B1#1/2 in LNCaP-AI and C4-2 cells. C-D, Colony formation and EdU assays were carried out to evaluate cell proliferation under HNRNPA2B1 silence. E-F, Cell apoptosis in two CRPC-like cells with or without HNRNPA2B1 suppression was detected via TUNEL assay and flow cytometry analysis. ** $P < 0.01$

for PCa, and proposed *C1QTNF3* as a promising gene for PCa utilizing RankProd method.

The nine hub genes selected in this study include *HNRNPA2B1*, *GOLGA8B*, *MAPK8IP3*, *CCDC14*, *MRPS25*, *RBM25*, *RBM33*, *TARDBP* and *FAM156B*. All these hub genes were oncogenes and tightly

related to the clinical characteristics of CRPC. Most of these key genes were identified as important biomarkers in some other diseases, such as breast cancer,¹⁷ lung cancer,¹⁸ pancreatic ductal adenocarcinoma,¹⁹ clear cell renal cell carcinoma²⁰ and so on. Eventually, *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* were picked

FIGURE 7 Genetic alterations associated with 9 hub genes. A, A visual summary of Genetic alterations (data from SU2C, MICH, SU2C219 and MPC Project) shows the genetic alteration of 9 hub genes. B, The total alteration frequency of 9 hub genes is illustrated. C-E, Expression difference among amplification, wild-type and deletion groups of HNRNPA2B1, *GOLGA8B* and *MAPK8IP3* in the TCGA-PRAD data set



out to be significantly associated with tumour prognosis. In order to further investigate potential mechanism of 9 hub genes, KEGG pathway analyses showed that genes in black module were mainly enriched in spliceosome pathway, RNA transport, surveillance and degradation pathways. Furthermore, GO analysis revealed that the hub genes were mainly associated with RNA splicing, mRNA processing, cilium morphogenesis and apoptotic process. Among the significantly enriched pathways, RNA transport and spliceosome pathway have been proven to be associated with tumour progression and prognosis in some cancer types, such as colorectal cancer,²¹ hepatocellular carcinoma,²² oesophageal squamous cell carcinoma²³ and so on.

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are RNA-binding proteins which are expressed in most human tissues. They stick to nascent transcript and are mainly involved in DNA repair, messenger RNA (mRNA) biogenesis, gene expression regulation, cell signalling and telomere biogenesis.²⁴ Previous studies have already investigated the prognostic role of *HNRNPA2B1* in many cancer types. Zhou et al²⁵ discovered that *HNRNPA2B1* was over-expressed in breast cancers, and the same conclusion could be drawn in lung cancers.^{18,26} Deng et al demonstrated that *HNRNPA2B1* was a driver oncogene in glioblastoma and acted as a predictor of overall survival.²⁷ Carles Barceló et al found that *HNRNPA2B1* interacted with oncogenic *KRAS* in pancreatic ductal adenocarcinoma cells.¹⁹ *GOLGA8B* is a protein-coding gene, which is supposed to be an important factor in the Golgi structure maintenance. The association between *GOLGA8B* and tumour progression or prognosis has not been explored deeply. The only corresponding study was reported by Wang et al, they found that over-expression of *GOLGA8B* might result in poor overall survival and *GOLGA8B* was a potential therapeutic target in patients with clear cell renal cell carcinoma.²⁰ *MAPK8IP3*, also known as c-Junamino-terminal kinase (*JNK*)/stress-activated protein kinase-associated protein 1 (*JSAP1*), was originally thought to be a binding protein for *JNK* and was suggested as a scaffold protein in neuronal cells via mammalian *JNK* signalling pathways.^{28,29} Furthermore, many studies have suggested the important role of *JSAP1* in certain neuronal processes, including axon elongation and branching.³⁰⁻³²

Mammalian mitochondrial ribosomal proteins are encoded by nuclear genes and are involved in the protein synthesis within the mitochondrion. Mitochondrial ribosomes (mitoribosomes) consist of a small 28S subunit and a large 39S subunit. *MRPS25*, one of 15 structural subunits of the 28S subunit, does not have a bacterial homolog.³³ Role of *MRPS25* has not been explored in CRPC previously. Study conducted by Enrico Bugiardini et al showed that *MRPS25* mutations could impair mitochondrial translation and cause encephalomyopathy.³⁴ *RBM25* and *RBM33* are family memberships of RNA-binding proteins (RBPs), which are involved in the co- and post-transcriptional process, influencing mRNA biogenesis, transport, stability, modification and cellular localization. However, rare research has been made on the role of *RBM25* and *RBM33*. In this study, over-expression of *RBM 25* and *RBM 33* was closely related

to CRPC phenotype. *FAM156B*, also known as *TMEM29B*, is a protein-coding gene with ubiquitous expression in skin, placenta, spleen, fat, adrenal and so on. Zané Lombard et al³⁵ found that *FAM156B* was a candidate gene utilizing a computational approach. Activation of HIV-1 gene expression by the transactivator Tat is dependent on an RNA regulatory element (TAR) located downstream of the transcription initiation site. *TARDBP*, a TAR DNA binding protein, is ubiquitous expressed, especially in lymph node and appendix. Protein encoded by *TARDBP* is a transcriptional repressor which binds to integrated TAR DNA chromosomally and represses HIV-1 transcription. There are few references on the role of *CCDC14* on transcriptome level. Huang et al³⁶ identified *CCDC14* as a candidate gene associated with autism spectrum disorder. Elif Nur Firat-Karalar et al reported that *CCDC14* played an important role in the interactions among centrosome components.³⁷

Recently, many previous studies were developed to identify key genes in CRPC by integrated bioinformatics methods. Wang et al³⁸ discovered *PTPRR* and *JAG1* as key genes which were tightly related to CRPC phenotype and long-term survival outcomes. Furthermore, Liang et al³⁹ investigated core genes for CRPC via bioinformatics analysis, and they found exisulind and phosphodiesterase-4 inhibitors were potential drugs for CRPC. Consistent with the finding revealed by Song et al,¹⁵ we also uncovered that the hub genes were associated with PCa progression. Nevertheless, there also existed differences between our present study and theirs. Song et al¹⁵ identified a module significantly related to Gleason score. In our study, we identified a module that could distinguish CRPC from HSPC and the hub genes were related to PCa prognosis and survival. And we aimed at finding out biomarkers for CRPC and validating their effectiveness.

Co-expression analysis is a useful technique for multigene analysis of large data sets. Based on these findings, we can apply the hub genes into clinical use in the future. However, some limitations in this study should be taken into consideration. First of all, it is a retrospective research, and data of this research were extracted from public database. Further large-cohort, multicentre and prospective studies are needed to validate these hub genes that may be related to tumour progression and prognosis. Secondly, this current study is lack of the in vivo experiments of investigating the molecular mechanisms to provide clinical guidance.

5 | CONCLUSION

Our study picked out and characterized 9 DEGs and significant gene module closely related to CRPC phenotype using WGCNA. Additionally, *HNRNPA2B1*, *GOLGA8B* and *MAPK8IP3* were identified to be tightly associated with tumour progression and prognosis. Further in vivo experiments are needed before clinical application.

ACKNOWLEDGEMENT

This work was supported by grants from the National Natural Science Foundation of China (No. 81702520), Medical Research Project of Jiangsu Provincial Health and Family Planning Commission

(No. H2018052), Research Project of Jiangsu Cancer Hospital (No. ZN201602), and the Young Talents Programme of Jiangsu Cancer Hospital (No. 2017YQL-04).

CONFLICT OF INTEREST

The authors declare no conflicts of interests.

AUTHOR CONTRIBUTION

Yifei Cheng: Conceptualization (equal); Data curation (lead); Formal analysis (equal); Software (equal); Writing-original draft (lead). **Lu Li:** Data curation (equal); Formal analysis (equal); Investigation (equal); Software (lead); Supervision (equal); Writing-original draft (supporting). **Zongshi Qin:** Formal analysis (equal); Investigation (equal); Methodology (lead); Supervision (equal); Validation (lead); Visualization (equal). **Xiao Li:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (lead); Methodology (equal); Project administration (lead); Visualization (equal); Writing-review & editing (supporting). **Feng Qi:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (supporting); Resources (equal); Software (equal); Validation (equal); Writing-review & editing (lead).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the TCGA data set (<https://portal.gdc.cancer.gov/>) and GEO data set (<http://www.ncbi.nlm.nih.gov/geo/>).

ORCID

Feng Qi  <https://orcid.org/0000-0003-1061-7058>

REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394-424.
- Huggins C, Hodges CV. Studies on prostatic cancer. I. The effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate. *J Urol.* 2002;167(2 Part 2):948-951.
- Yuan X, Balk SP. Mechanisms mediating androgen receptor reactivation after castration. *Urol Oncol.* 2009;27:36-41.
- Attard G, Cooper CS, de Bono JS. Steroid hormone receptors in prostate cancer: a hard habit to break? *Cancer Cell.* 2009;16:458-462.
- Jiao L, Deng Z, Xu C, et al. miR-663 induces castration-resistant prostate cancer transformation and predicts clinical recurrence. *J Cell Physiol.* 2014;229:834-844.
- Inoue T, Segawa T, Kamba T, et al. Prevalence of skeletal complications and their impact on survival of hormone refractory prostate cancer patients in Japan. *Urology.* 2009;73:1104-1109.
- Smith MR, Kabbavar F, Saad F, et al. Natural history of rising serum prostate-specific antigen in men with castrate nonmetastatic prostate cancer. *J Clin Oncol.* 2005;23:2918-2925.
- Kirby M, Hirst C, Crawford ED. Characterising the castration-resistant prostate cancer population: a systematic review. *Int J Clin Pract.* 2011;65:1180-1192.
- Antonarakis ES, Eisenberger MA. Expanding treatment options for metastatic prostate cancer. *N Engl J Med.* 2011;364:2055-2058.
- Wu G, Huang S, Nastiuk KL, et al. Variant allele of HSD3B1 increases progression to castration-resistant prostate cancer. *Prostate.* 2015;75:777-782.
- Russo JW, Gao C, Bhasin SS, et al. Downregulation of dipeptidyl peptidase 4 accelerates progression to castration-resistant prostate cancer. *Cancer Res.* 2018;78:6354-6362.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 2008;9:559.
- Tang J, Kong D, Cui Q, et al. Prognostic genes of breast cancer identified by gene co-expression network analysis. *Front Oncol.* 2018;8:374.
- Barbieri CE, Chinnaiyan AM, Lerner SP, Swanton C, Rubin MA. The emergence of precision urologic oncology: a collaborative review on biomarker-driven therapeutics. *Eur Urol.* 2017;71:237-246.
- Song ZY, Chao F, Zhuo Z, Ma Z, Li W, Chen G. Identification of hub genes in prostate cancer using robust rank aggregation and weighted gene co-expression network analysis. *Aging (Albany NY).* 2019;11:4736-4756.
- Hou Q, Bing ZT, Hu C, et al. RankProd combined with genetic algorithm optimized artificial neural network establishes a diagnostic and prognostic prediction model that revealed C1QTNF3 as a biomarker for prostate cancer. *EBioMedicine.* 2018;32:234-244.
- Hu Y, Sun Z, Deng J, et al. Splicing factor hnRNP2B1 contributes to tumorigenic potential of breast cancer cells through STAT3 and ERK1/2 signaling pathway. *Tumour Biol.* 2017;39:1010428317694318.
- Dowling P, Pollard D, Larkin A, et al. Abnormal levels of heterogeneous nuclear ribonucleoprotein A2B1 (hnRNP2B1) in tumour tissue and blood samples from patients diagnosed with lung cancer. *Mol Biosyst.* 2015;11:743-752.
- Barcelo C, Etchin J, Mansour MR, et al. Ribonucleoprotein HNRNP2B1 interacts with and regulates oncogenic KRAS in pancreatic ductal adenocarcinoma cells. *Gastroenterology.* 2014;147:882-892.e8.
- Wang J, Zhang C, He W, Gou X. Construction and comprehensive analysis of dysregulated long non-coding RNA-associated competing endogenous RNA network in clear cell renal cell carcinoma. *J Cell Biochem.* 2019;120(2):2576-2593.
- Lopez-Sanchez LM, Jimenez-Izquierdo R, Penarando J, et al. SWATH-based proteomics reveals processes associated with immune evasion and metastasis in poor prognosis colorectal tumours. *J Cell Mol Med.* 2019;23(12):8219-8232.
- Iguchi T, Komatsu H, Masuda T, et al. Increased copy number of the gene encoding SF3B4 indicates poor prognosis in hepatocellular carcinoma. *Anticancer Res.* 2016;36:2139-2144.
- He W, Chen L, Yuan K, Zhou Q, Peng L, Han Y. Gene set enrichment analysis and meta-analysis to identify six key genes regulating and controlling the prognosis of esophageal squamous cell carcinoma. *J Thorac Dis.* 2018;10:5714-5726.
- Han SP, Tang YH, Smith R. Functional diversity of the hnRNPs: past, present and perspectives. *Biochem J.* 2010;430:379-392.
- Zhou J, Allred DC, Avis I, et al. Differential expression of the early lung cancer detection marker, heterogeneous nuclear ribonucleoprotein-A2/B1 (hnRNP-A2/B1) in normal breast and neoplastic breast cancer. *Breast Cancer Res Treat.* 2001;66:217-224.
- Fielding P, Turnbull L, Prime W, Walshaw M, Field JK. Heterogeneous nuclear ribonucleoprotein A2/B1 up-regulation in bronchial lavage specimens: a clinical marker of early lung cancer detection. *Clin Cancer Res.* 1999;5:4048-4052.
- Deng J, Chen S, Wang F, et al. Effects of hnRNP A2/B1 knockdown on inhibition of glioblastoma cell invasion, growth and survival. *Mol Neurobiol.* 2016;53:1132-1144.

28. Kelkar N, Gupta S, Dickens M, Davis RJ. Interaction of a mitogen-activated protein kinase signaling module with the neuronal protein JIP3. *Mol Cell Biol*. 2000;20:1030-1043.
29. Ito M, Yoshioka K, Akechi M, et al. JSAP1, a novel jun N-terminal protein kinase (JNK)-binding protein that functions as a Scaffold factor in the JNK signaling pathway. *Mol Cell Biol*. 1999;19:7539-7548.
30. Sun T, Yu N, Zhai LK, et al. c-Jun NH2-terminal kinase (JNK)-interacting protein-3 (JIP3) regulates neuronal axon elongation in a kinesin- and JNK-dependent manner. *J Biol Chem*. 2013;288:14531-14543.
31. Bilimoria PM, de la Torre-Ubieta L, Ikeuchi Y, Becker EB, Reiner O, Bonni A. A JIP3-regulated GSK3beta/DCX signaling pathway restricts axon branching. *J Neurosci*. 2010;30:16766-16776.
32. Suzuki A, Arikawa C, Kuwahara Y, et al. The scaffold protein JIP3 functions as a downstream effector of the small GTPase ARF6 to regulate neurite morphogenesis of cortical neurons. *FEBS Lett*. 2010;584:2801-2806.
33. Cavdar Koc E, Burkhart W, Blackburn K, Moseley A, Spremulli LL. The small subunit of the mammalian mitochondrial ribosome. Identification of the full complement of ribosomal proteins present. *J Biol Chem*. 2001;276:19363-19374.
34. Bugiardini E, Mitchell AL, Rosa ID, et al. MRPS25 mutations impair mitochondrial translation and cause encephalomyopathy. *Hum Mol Genet*. 2019;28:2711-2719.
35. Lombard Z, Park C, Makova KD, Ramsay M. A computational approach to candidate gene prioritization for X-linked mental retardation using annotation-based binary filtering and motif-based linear discriminatory analysis. *Biol Direct*. 2011;6:30.
36. Huang H, Cheng S, Ding M, et al. Integrative analysis of transcriptome-wide association study and mRNA expression profiles identifies candidate genes associated with autism spectrum disorders. *Autism Res*. 2019;12:33-38.
37. Firat-Karalar EN, Rauniyar N, Yates JR 3rd, Stearns T. Proximity interactions among centrosome components identify regulators of centriole duplication. *Curr Biol*. 2014;24:664-670.
38. Wang JL, Wang Y, Ren GP. Identification of PTPRR and JAG1 as key genes in castration-resistant prostate cancer by integrated bioinformatics methods(). *J Zhejiang Univ Sci B*. 2020;21:246-255.
39. Liang X, Hu K, Li D, et al. Identification of core genes and potential drugs for castration-resistant prostate cancer based on bioinformatics analysis. *DNA Cell Biol*. 2020;39(5):836-847.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Cheng Y, Li L, Qin Z, Li X, Qi F. Identification of castration-resistant prostate cancer-related hub genes using weighted gene co-expression network analysis. *J Cell Mol Med*. 2020;24:8006–8017. <https://doi.org/10.1111/jcmm.15432>