

This article is licensed under a Creative Commons Attribution-NonCommercial NoDerivatives 4.0 International License.

## Correlating Transcriptional Networks to Papillary Renal Cell Carcinoma Survival: A Large-Scale Coexpression Analysis and Clinical Validation

Xingliang Feng,<sup>\*1</sup> Meng Zhang,<sup>\*†1</sup> Jialin Meng,<sup>\*</sup> Yongqiang Wang,<sup>†</sup> Yi Liu,<sup>\*</sup> Chaozhao Liang,<sup>\*</sup> and Song Fan<sup>\*</sup>

<sup>\*</sup>Department of Urology, The First Affiliated Hospital of Anhui Medical University, Institute of Urology, Anhui Province Key Laboratory of Genitourinary Diseases, Anhui Medical University, Hefei, China

<sup>†</sup>Urology Institute of Shenzhen University, The Third Affiliated Hospital of Shenzhen University, Shenzhen University, Shenzhen, China

We aimed to investigate the potential mechanisms of progression and identify novel prognosis-related biomarkers for papillary renal cell carcinoma (PRCC) patients. The related data were derived from The Cancer Genome Atlas (TCGA) and then analyzed by weighted gene coexpression network analysis (WGCNA). The correlation between each module and the clinical traits were analyzed by Pearson's correlation analysis. Pathway analysis was conducted to reveal potential mechanisms. Hub genes within each module were screened by intramodule analysis, and visualized by Cytoscape software. Furthermore, important hub genes were validated in an external dataset and clinical samples. A total of 5,839 differentially expressed genes were identified. By using WGCNA, we identified 21 core regulatory gene clusters based on 289 PRCC samples. We found many modules were significantly associated with clinicopathological characteristics. The gray, pink, light yellow, and salmon modules served as prognosis indicators for PRCC patients. Pathway enrichment analyses found that the hub genes were significantly enriched in the cancer-related pathways. With the external Gene Expression Omnibus (GEO) validation dataset, we found that PCDH12, GPR4, and KIF18A in the pink and yellow modules were continually associated with the survival status of PRCC, and their expressions were positively correlated with pathological grade. Notably, we randomly chose PCDH12 for validation, and the results suggested that the PRCC patients with higher pathological grades (II + III) mostly had higher PCDH12 protein expression levels compared with those patients in grade I. These validated hub genes play critical roles in the prognosis prediction of PRCC and serve as potential biomarkers for future personalized treatment.

**Key words:** Papillary renal cell carcinoma (PRCC); Weighted gene coexpression network analysis (WGCNA); Hub gene; Prognosis

### INTRODUCTION

Renal cell carcinoma (RCC) is regarded as one of the most serious human diseases. In the US, there were an estimated 73,820 (~4.18%) new cancer cases and 14,770 (~2.43%) premature deaths in 2019<sup>1</sup>. The disease encompasses several major subtypes, including clear cell RCC (ccRCC; 70%–80%), papillary RCC (PRCC; 10%–15%), chromophobe RCC (chRCC; 3%–5%), and renal oncocytoma (RO). Each subtype has different genetic characteristics and histological features<sup>2–6</sup>. PRCC is characterized by the presence of a fibrous vascular core of tumor cells arranged in a

papillary pattern<sup>7</sup>, and there is no specific treatment option available for PCRR patients. Few studies investigate the underlying mechanisms of PRCC, delaying the development of new drugs. Besides, for metastatic PRCC patients, they commonly have an unfavorable prognosis, particularly compared to patients with ccRCC<sup>8</sup>. Therefore, it is necessary to identify new molecular markers that can be used to predict disease stage and clinical outcome in patients with PRCC. This can help understand its pathological mechanisms and provide personalized treatment.

To date, many biomarkers for PRCC have been identified, including several markers that can predict

<sup>1</sup>These authors provided equal contribution to this work.

Address correspondence to Prof. Song Fan, Department of Urology, The First Affiliated Hospital of Anhui Medical University, Institute of Urology, Anhui Province Key Laboratory of Genitourinary Diseases, Anhui Medical University, Jixi Road, Shushan District, Hefei, Anhui 230022, China. Tel: 86-0551-62923861; E-mail: [songfandoctor@gmail.com](mailto:songfandoctor@gmail.com)

treatment outcomes and clinical prognoses, such as *VHL*, *VEGF*, *CAIX*, and *HIF1 $\alpha$ /2 $\alpha$*  mutations<sup>9</sup>. The mutation rates of these genes are low, and we still lack sufficient gene markers to predict prognosis for PRCC patients<sup>10</sup>. The rapid breakthroughs in genome-wide sequencing technology provide new ideas for the study of various clinical problems and pathological mechanisms related to cancer<sup>11</sup>. The Cancer Genome Atlas (TCGA) can help to improve diagnostic methods and ultimately improve the survival and prognosis of cancer patients. TCGA has generated an extensive collection of clinical information and genetic sequencing data that allows the systematic analysis of the underlying molecular mechanisms involved in the various clinical characteristics associated with cancer, such as tumor grade, pathological stage, histological type, diagnosis, and prognosis<sup>12</sup>. Weighted gene coexpression network analysis (WGCNA) explores the relationship between different genomes and clinical features by constructing a free-scale gene coexpression network<sup>13</sup>. The WGCNA algorithm is widely used to screen biological processes and therapeutic targets related to cancer as well as specific biomarkers associated with complex diseases<sup>14–16</sup>. Similarly, WGCNA is typically used to identify critical genes that are significantly associated with clinical indicators of tumor progression, including the stage, grade, and metastasis of different types of tumors<sup>17,18</sup>.

In this study, by generating a coexpression network analysis, we identified differences in expression between different subclassification and functional modules related to the prognosis of patients with PRCC. The flowchart of our work is illustrated in Figure 1. Additionally, the hub gene analysis may have important clinical significance and allow us to identify critical genes in each module that can be used as prognostic biomarkers or therapeutic targets for PRCC patients.

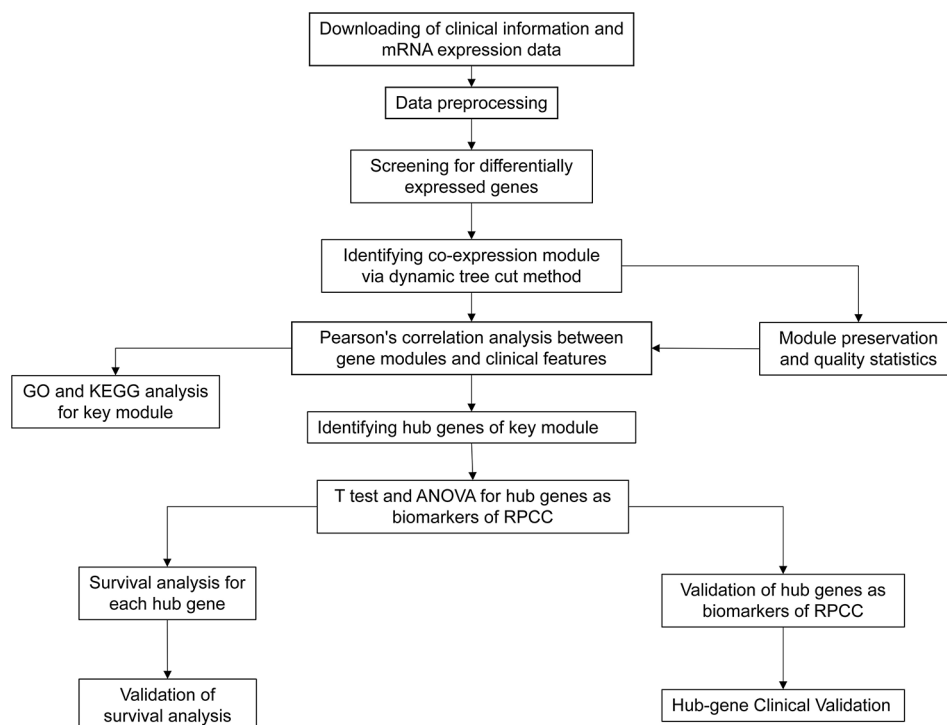
## MATERIALS AND METHODS

### Data Acquirements

All PRCC patients were recruited from the TCGA project ([https://xenabrowser.net/datapages/?cohort=TCGA%20Kidney%20Papillary%20Cell%20Carcinoma%20\(KIRP\)&removeHub=https%3A%2F%2Fxcna.treehouse.gi.ucsc.edu%3A443](https://xenabrowser.net/datapages/?cohort=TCGA%20Kidney%20Papillary%20Cell%20Carcinoma%20(KIRP)&removeHub=https%3A%2F%2Fxcna.treehouse.gi.ucsc.edu%3A443)), which is a cohort study of PRCC<sup>19</sup>. In the current study, a cohort of 289 PRCC patients and 32 normal controls with matched mRNA expression profiles were used along with their clinical data.

### Identification of Differentially Expressed Genes (DEGs) and Gene Set Enrichment Analysis (GSEA) Analysis

An in-depth analysis of DEGs between 289 tumor specimens and 32 normal controls was performed using the R package DESeq2 v1.20.0<sup>20</sup>. Significant differentially



**Figure 1.** Flow diagram of strategy for data preparation, processing, analysis, and validation used in this study.

expressed genes were selected with both  $p_{\text{adj}} < 0.05$  and  $\log_2(\text{fold change}) \geq 1$ . In addition, GSEA analysis was conducted based on gseapy v0.9.3, which could also display the gene distribution differences between the PRCC patients and normal controls.

#### WGCNA and Pathway Enrichment

Gene coexpression network analysis is a systems biology method and was performed by the WGCNA package<sup>21</sup>. In the coexpression network, the DEGs are represented by nodes, and the connectivity between the genes is defined by the correlation of the gene expression patterns<sup>22</sup>. Analysis of the network topology was used to determine the soft threshold power. The soft threshold power was used to successively calculate the gene coexpression similarity and the adjacency. The adjacency was then transformed into a topological overlap matrix (TOM). The eigengene is defined using the singular value decomposition (SVD) of the module expression matrix<sup>23,24</sup>.  $X^{(l)}$  ( $\chi_{il}^{(l)}$ ) represents the gene expression matrix of the  $l$ th module. The module gene corresponds to the indexes  $i = 1, 2, \dots, n_l$ , and the ID of microarray samples corresponds to the indexes  $l = 1, 2, \dots, m$ . The SVD of  $X^{(l)}$  is expressed by the following formula:

$$X^{(l)} = U\Sigma V^T \quad (1)$$

The columns of orthogonal matrices  $U$  is called a left singular vectors, and the columns of orthogonal matrices  $V$  is called a right singular vectors.  $\Sigma$  value is only in the main diagonal, and we call it as singular values, while other elements are all equal to zero.

More precisely,  $U^{(l)}$  represents a  $n^{(l)} \times m$  matrix with orthonormal columns;  $V^{(l)}$  represents a  $m \times m$  orthogonal matrix, and  $\Sigma^{(l)}$  represents the  $m \times m$  diagonal matrix of the singular values  $\{|d_i^{(l)}|\}$ .

The matrices  $V^{(l)}$  and  $\Sigma^{(l)}$  are denoted by

$$\begin{aligned} V^{(l)} &= (v_1^{(l)} v_2^{(l)} \dots v_m^{(l)}), \\ \Sigma^{(l)} &= \text{diag}\{|d_1^{(l)}|, |d_2^{(l)}|, \dots, |d_m^{(l)}|\}. \end{aligned} \quad (2)$$

Assuming that the singular values  $|d_i^{(l)}|$  are arranged in a nonascending order, we call the first column of  $V^{(l)}$  as the module eigengene:

$$E^{(l)} = V_1^{(l)} \quad (3)$$

Since the direction of each singular vector is undefined, we determine the direction of the characteristic gene by constraining the direction of each characteristic gene to make it positively correlated with the average gene expression of each module gene<sup>24</sup>. Finally, using the TOM for hierarchical clustering, the dynamic tree cut algorithm was applied to the module screening, and then the gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed

on the gene module, after which the modules related to PRCC were characterized.

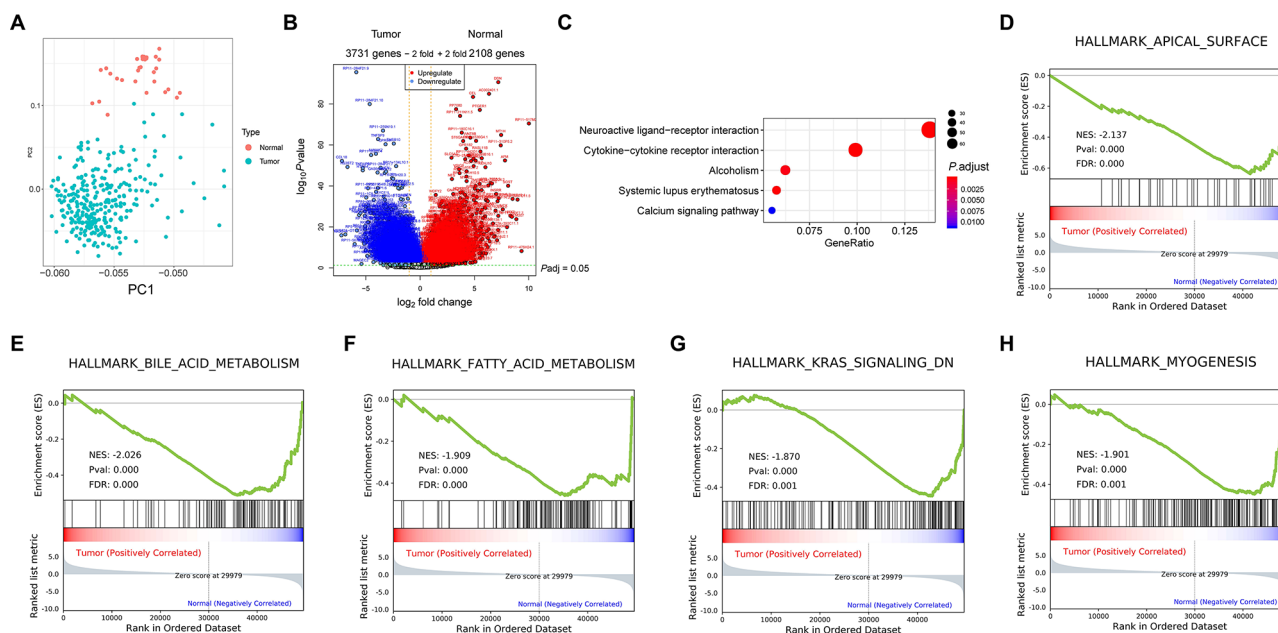
#### Survival Analysis and Hub Gene Validation

We used the “survival”<sup>25</sup> and “survminer”<sup>26</sup> R packages to calculate the correlation between each module and overall survival (OS) and recurrence-free survival (RFS), with a value of  $p < 0.05$  considered as statistically significant. In addition, the connections between the modules and the clinicopathological features were determined by Pearson’s correlation analysis, and a value of  $p < 0.05$  was treated as statistically significant. The gene/eigengene expression differences in different subgroups were determined by  $t$ -test (two groups) or one-way analysis of variance (ANOVA) analysis (more than two groups). A hub gene is one of a series of genes that determine the characteristics of a module, and it has the highest degree of connectivity in a gene module. Based on their importance, we chose the top 10 hub genes within the significant modules and used an additional GEO dataset (GSE2748) to validate their usage for the prediction of clinical characteristics, such as prognosis and tumor grade, and confirmed their role in cellular function in vitro and in prognosis prediction by immunohistochemistry (IHC). In addition, the detail steps of the IHC were demonstrated in our previously published work<sup>27</sup>. The  $t$ -test (unpaired) was used to determine the protein expression differences between low- and high-grade PRCC patients. The PCDH12 antibody (catalog # PA5-20703) was purchased from Invitrogen (Carlsbad, CA, USA). And the IHC assay was performed on a tissue array containing 37 PRCC cases, which was bought from Xi’an AiDi Biotechnology Ltd. Co. (Xi’an, Shanxi Province, China). All patients gave written informed consent for the use of the tissue samples for research purposes. For the patients enrolled in the TCGA project, they have announced the ethics and policies on their own website (<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/history/policies>). The research contents and research programs were reviewed and approved by the Ethics Committee of the First Affiliated Hospital of Anhui Medical University (anyiyifuyuanlunshen-kuai-PJ-2019-09-11).

## RESULTS

### DEGs Screening and GSEA Analyses

The expression matrices for 289 PRCC samples and 32 normal controls were obtained from the TCGA database. Based on thresholds defined by  $\log_2(\text{fold change}) \geq 1$  and a  $p_{\text{adj}} < 0.05$ , 5,839 DEGs were screened in the PRCC and normal control groups, of which 2,108 were upregulated and 3,371 were downregulated (Fig. 2A and B). Table S1 contains all the differential genes used in Figure 2B (see supplemental Table S1, available at <https://pan.baidu.com>).



**Figure 2.** Differentially expressed genes and pathway enrichment analyses. (A) Principal component analysis (PCA) for the papillary renal cell carcinoma (PRCC) tumor and normal tissues. (B) The differentially expressed genes between PRCC tumor and normal tissues ( $p < 0.05$ ). (C) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment of the differentially expressed genes ( $p < 0.05$ ). (D–H) Gene Set Enrichment Analysis (GSEA) analyses between PRCC tumor and normal tissues ( $p < 0.05$ ).

com/s/1xRGQBJHNJrQnwpWfrEQUPQ). We performed KEGG enrichment for these DEGs and found that these genes were enriched mostly within the cytokine–cytokine receptor interaction, neuroactive ligand–receptor interaction, and alcoholism pathways ( $p_{\text{adj}} < 0.05$ ) (Fig. 2C). These DEGs were then selected for subsequent analysis. In addition, we also conducted an online GSEA analysis, and the results suggested that genes in the apical surface, bile acid metabolism, fatty acid metabolism, KRAS signaling, and myogenesis pathways were significantly differentially expressed between PRCC cancer patients and normal controls ( $p_{\text{adj}} < 0.05$ ) (Fig. 2D–H).

#### Coexpression Network Construction

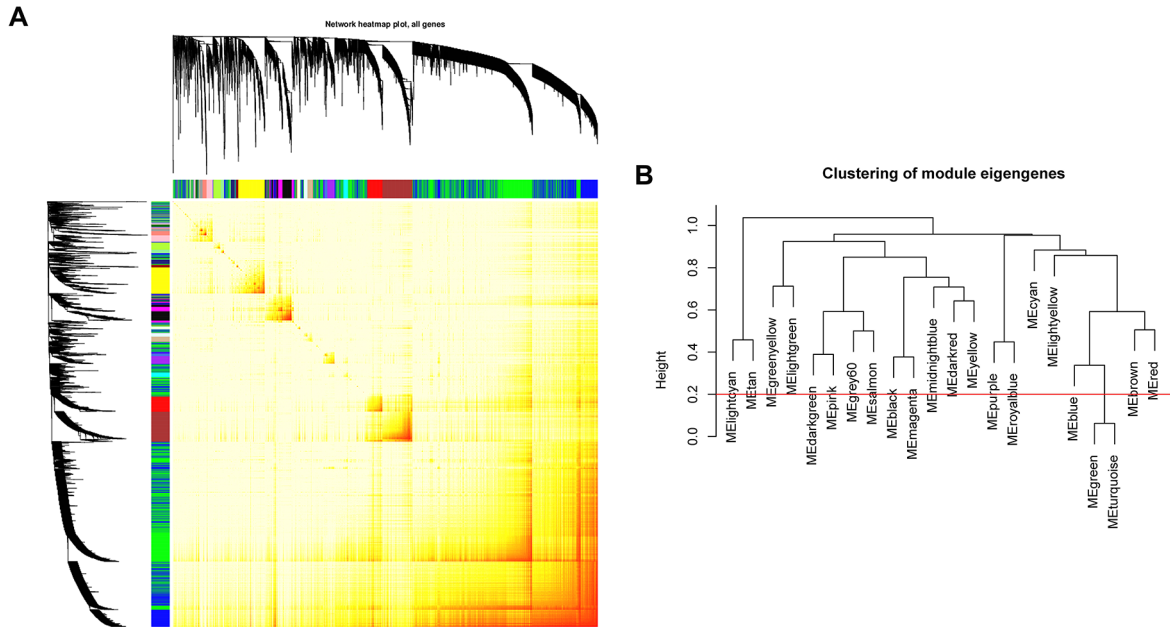
WGCNA was performed on 5,839 DEGs from 289 samples ( $p_{\text{adj}} < 0.05$ ). The intergene connectivity in the gene network attained a scale-free network distribution (Fig. 3A). Therefore, we calculated the eigengene for entire modules and clustered them based on their correlation to further quantify their coexpression similarity. After merging two similar modules, a total of 21 modules were obtained (eigengene of the module, height  $> 0.2$ ) (Fig. 3B).

#### Finding Modules of Clinical Significance and Identifying the Hub Genes of Modules

The identification of the gene modules most significantly associated with different clinical characteristics

has important biological implications, and the correlation between each module and the clinical features was observed in the module characteristic diagram (Fig. 4A), in which the largest correlation coefficient ( $r^2 = 0.5$ ,  $p = 2 \times 10^{-18}$ ) was observed between the pink module and the initial weights of the samples. In addition, we also observed that several modules were significantly associated with smoking (number of years) (negative: light cyan, purple, black, magenta, and light green modules; positive: green module) and new tumor events (negative: pink, salmon, midnight blue, and yellow modules) ( $p < 0.05$ ). It can be inferred that these modules were of clinical significance. Interestingly, we found several modules significantly associated gender differences, including the tan, royal blue, blue, dark green, pink, gray, midnight blue, dark red, yellow, yellow-green, and light green modules ( $p < 0.05$ ). Current work also analyzed the connection between sex and module eigengene expression, and the results are presented in Figure 4B. We then analyzed the correlations between these modules and clinicopathological features and found that the pink and yellow modules were significantly associated with the PRCC stage ( $p < 0.05$ ) (Fig. 4C) and could serve as markers for predicting the tumor stage. In addition, we also found that the light cyan, light yellow, pink, gray, salmon, dark red, yellow, yellow-green, and light green modules were associated with new tumor events ( $p < 0.05$ ) (Fig. 4D), which is partly consistent with the module characteristic diagram.





**Figure 3.** Visualization of the gene network using a Heatmap plot. (A) The heatmap depicts the topological overlap matrix (TOM) showing all genes used in the analysis. The light red color represents a low overlap, and the progressively darker red color represents higher overlap. (B) Hierarchical eigengene diagram of all samples.

### Prognostic Analyses and Network Construction

Pathological staging is an important factor that determines the prognosis of patients with PRCC. The higher the pathological staging, the worse the prognosis. The molecular biomarkers we screened for their correlation with pathological staging are also prognostic genes, in theory. In our work, we found that the gray, pink, light yellow, yellow, and salmon modules were significantly correlated with OS and RFS in PRCC ( $p < 0.05$ ) (Fig. 5). When this was combined with the results shown in Figures 4 and 5, we found that the pink and yellow modules had the highest clinical significance. Therefore, we used Cytoscape software to visualize these two modules (Fig. 6). Consistent with the hub gene analysis results, we found that *NOVA2*, *GPR4*, *MYCT1*, and *NOTCH4*, which were the hub genes in the yellow module, were located at the center of the network and were significantly associated with other genes. Similar results were obtained for the pink module.

### Pathway Enrichment and Independent Dataset Validation

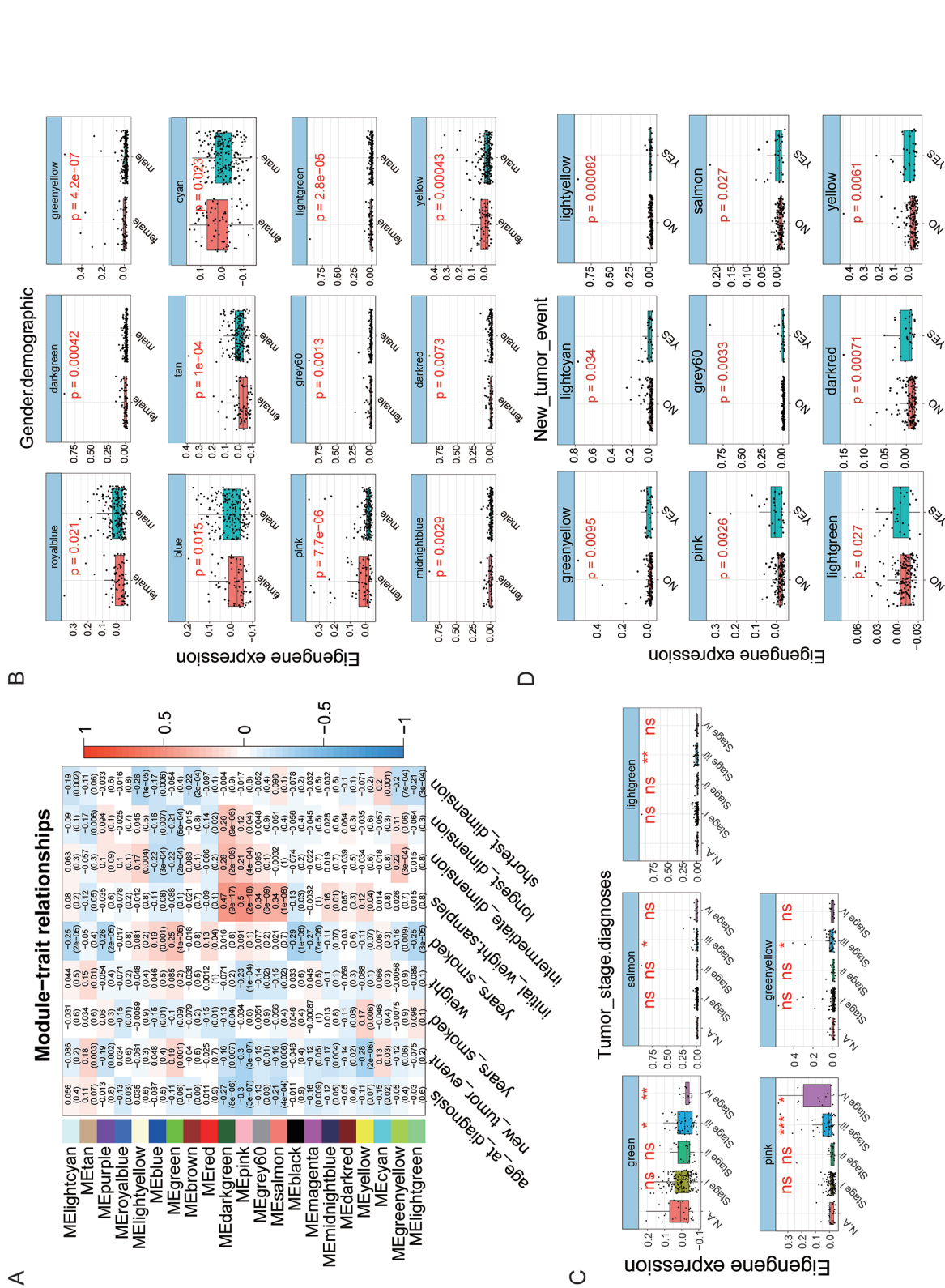
By pathway enrichment analysis, we identified the significant KEGG pathways and GO terms that were enriched in the significant modules when the Benjamin-adjusted  $p$  value threshold was  $< 0.05$ . As shown in Figure 7, for the gray 60 module, the module genes that were involved in the systemic lupus erythematosus, maturity onset diabetes of the young, alcoholism, chromosome

maintenance, DNA methylation, signaling by nuclear receptors, and epigenetic regulation of gene expression event pathways were significantly enriched ( $p_{\text{adj}} < 0.05$ ).

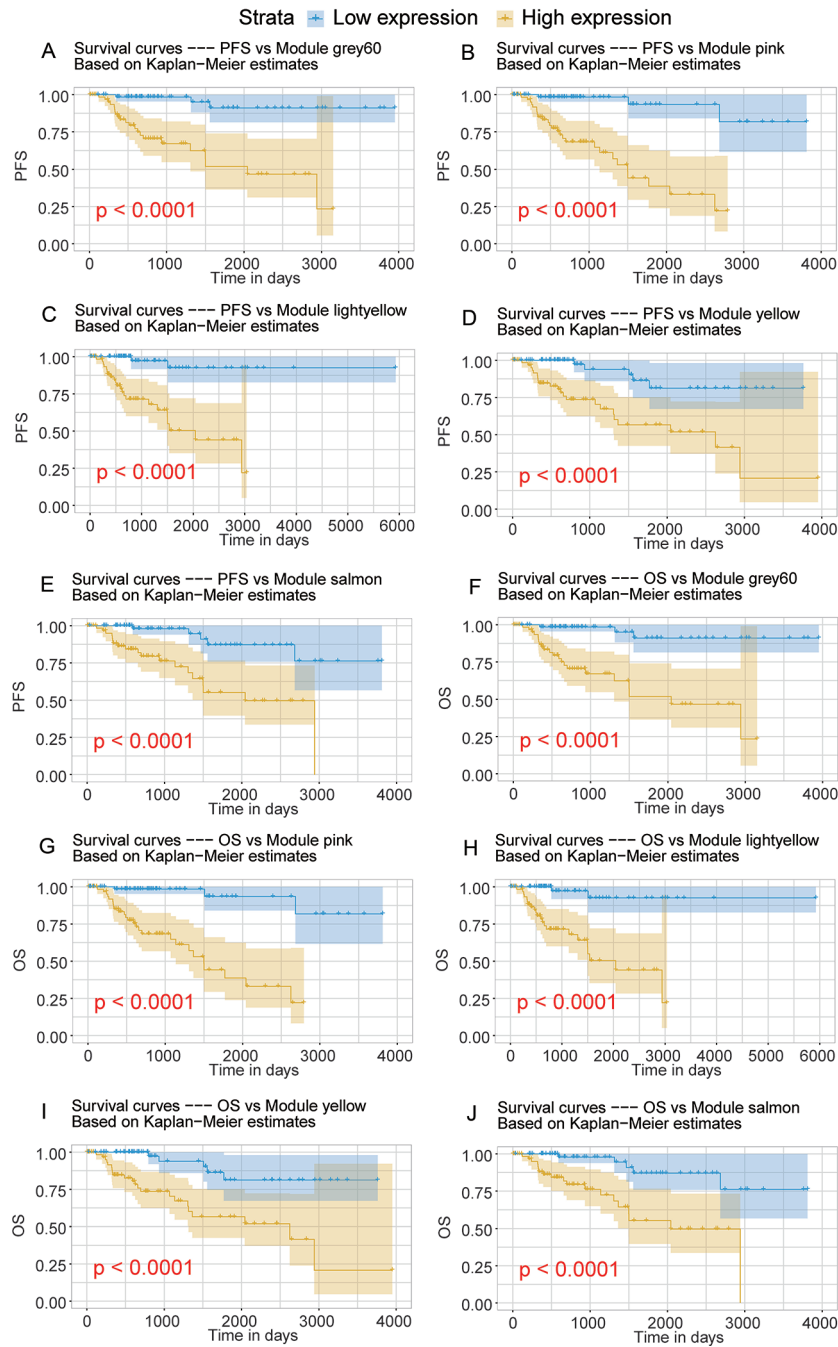
In addition, based on the importance of the pink and yellow modules, we chose the top 10 genes ( $p < 0.05$ , excluding those genes that were not associated with OS or RFS) and used additional GEO datasets to further validate their importance. As shown in Figure 8, we found that *PCDH12*, *GPR4*, and *KIF18A* were continually associated with survival status in PRCC ( $p < 0.05$ ) (Fig. 8A–C). Consistently, we also found that *KIF18A*, *GPR4*, and *PCDH12* were significantly positively associated with histological classification type 2B ( $p < 0.05$ ) (Fig. 8D), molecular classification class 2 (Fig. 8E), and grades 3 and 4 and were more highly expressed in tumor node metastasis stages 3 and 4 ( $p < 0.05$ ) (Fig. 8G) and the M1 stage ( $p < 0.05$ , *KIF18A* and *GPR4* only) (Fig. 8H). Based on the importance of these three genes, further IHC assay was performed on the PRCC clinical samples (Table 1). Moreover, we randomly chose *PCDH12* for IHC validation, and the results suggested that the PRCC patients with higher pathological grades (II + III) mostly had high *PCDH12* protein expression ( $p < 0.05$ ), a result consistent with our above findings (Fig. 9).

## DISCUSSION

It is extremely important to investigate the molecular markers used for the diagnosis and treatment of PRCC patients. However, there is a lack of sufficient clinical use



**Figure 4.** Correlation between module eigengene expression and clinicopathological features. (A) Module-trait relationships. (B) Module eigengene expression difference between males and females. (C) Module eigengene expression difference between different tumor stages. (D) Correlation between module eigengene expression and recurrent status. A value of  $p < 0.05$  was regarded as statistically significant.

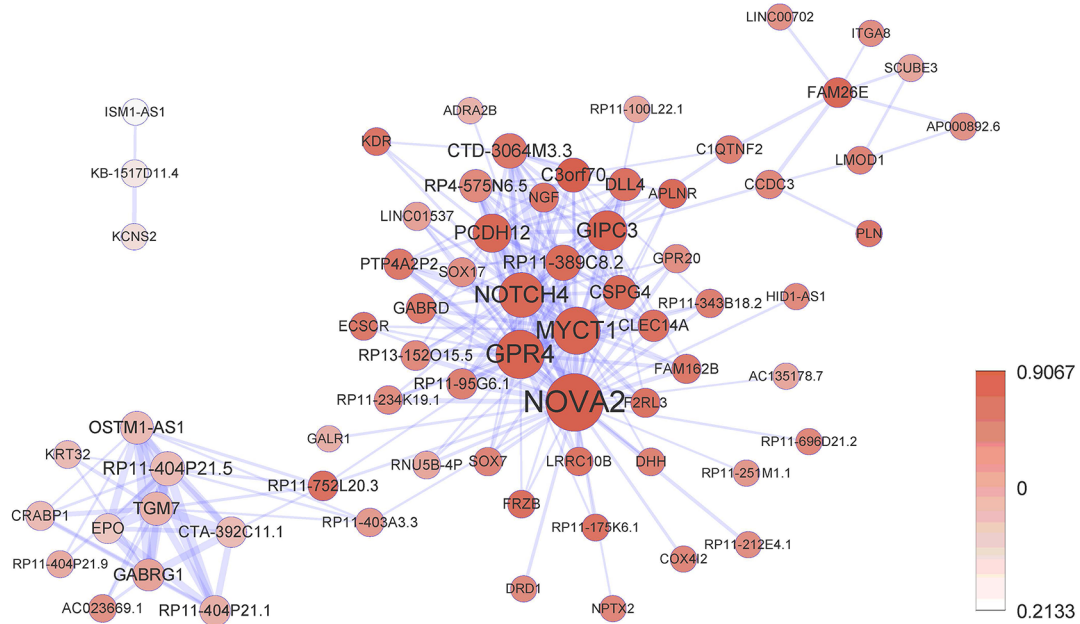


**Figure 5.** Survival analyses. (A–E) Correlation between the module eigengene expression and recurrence-free survival of PRCC patients. (F–J) Correlation between the module eigengene expression and overall survival of PRCC patients. A value of  $p < 0.05$  was regarded as statistically significant.

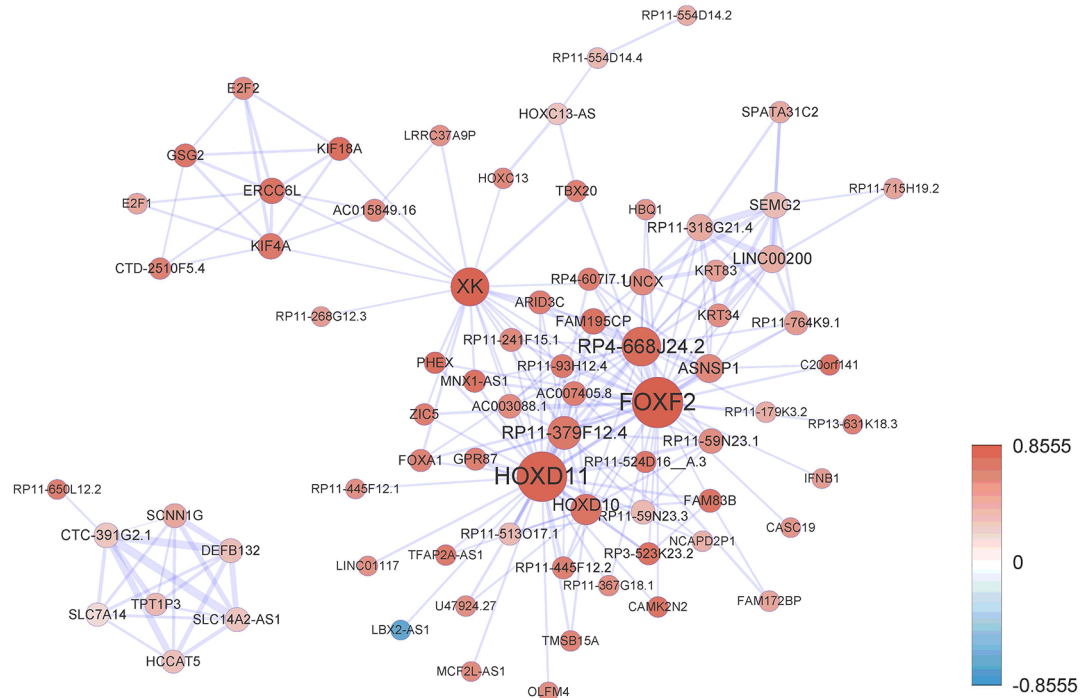
for the molecular markers currently associated with histological grade in PRCC. In this study, we used WGCNA to investigate PRCC-related coexpression modules and their clinical features. Significant modules, including the gray, pink, light yellow, yellow, and salmon modules, were shown to be correlated with OS and RFS in PRCC. The gray module was mainly related to pathways

involved in alcoholism, necroptosis, and DNA complex packaging. Previous studies have reported that ethanol, which is the main component of alcoholic beverages, is metabolized into acetaldehyde, which is classified as carcinogenic to humans<sup>28</sup>. Alcohol consumption has now been identified as a risk factor for human cancer, but there is still a lack of direct evidence for PRCC. In

A



B

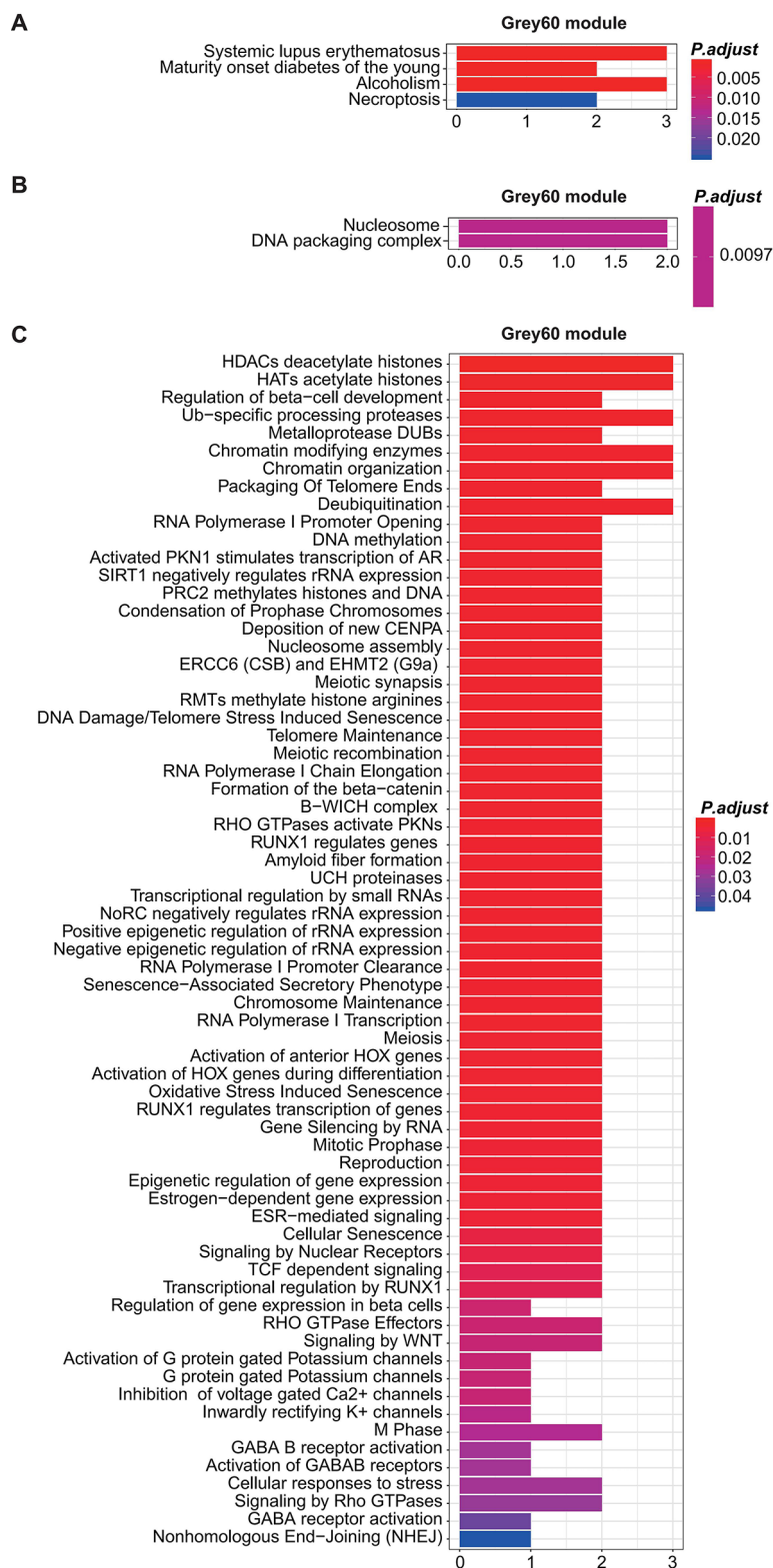


**Figure 6.** Network analyses for the hub genes. (A) Network analysis for the hub genes in yellow module. (B) Network analysis for the hub genes in the pink module.

addition, recent studies have shown that alcohol increases the number of cancer stem cells (CSCs) in cancer, which may be the basis of alcohol-induced tumors<sup>29-31</sup>. In addition, necroptosis involves the caspase-independent regulation of cell death that is executed by receptor-interacting

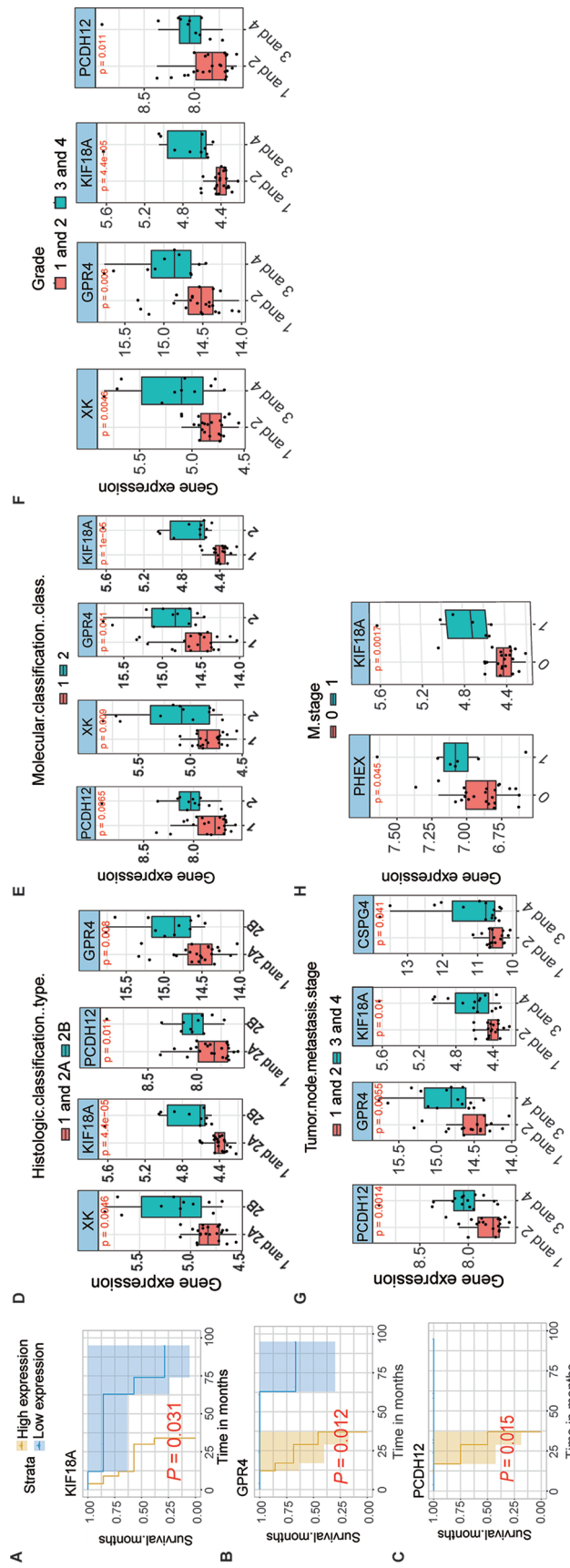
protein 1 (RIP1), receptor-interacting protein 3 (RIP3), and mixed lineage kinase domain-like protein (MLKL)<sup>32,33</sup>. In recent years, tumor treatment based on tumor necrosis has been considered as a new antitumor treatment strategy<sup>34,35</sup>.





**Figure 7.** KEGG analyses for the hub genes. The KEGG analyses for the hub genes in the gray 60 module. (A) KEGG analysis. (B) GO\_cellular component analysis. (C) Reactome analysis. A value of  $p < 0.05$  was regarded as statistically significant.





**Figure 8.** External validation of the key hub genes within yellow and pink modules. (A–C) Survival analyses for the key hub genes in the yellow and pink modules. (D) Correlation between hub gene expression and histologic classification. (E) Correlation between hub gene expression and molecular classification. (F) Correlation between hub gene expression and pathological grade. (G) Correlation between hub gene expression and tumor–node–metastasis stage. (H) Correlation between hub gene expression and metastasis stage. A value of  $p < 0.05$  was regarded as statistically significant.

**Table 1.** Clinical Characteristics of 37 Papillary Renal Cell Carcinoma (PRCC) Patients

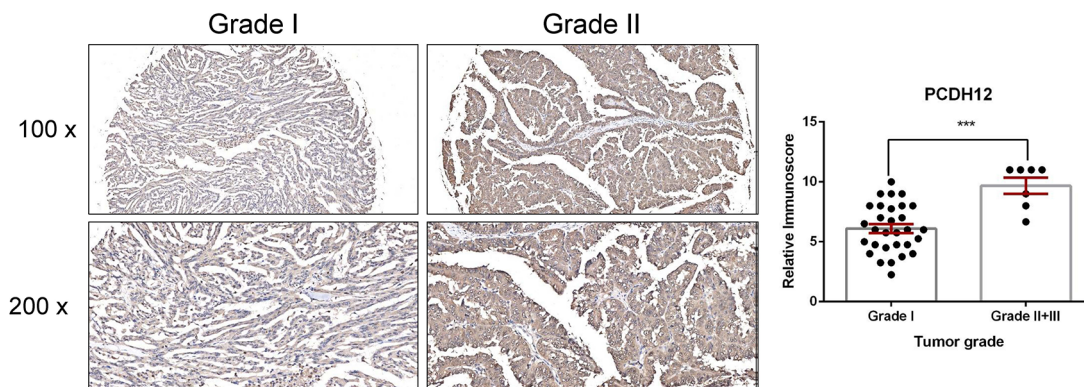
Characteristic	<i>n</i> (%)
<b>Gender</b>	
Male	24 (37.84)
Female	13 (35.14)
<b>Age (years)</b>	
≤55	23 (62.16)
>55	14 (37.84)
<b>Subtype</b>	
I	10 (27.03)
II	27 (72.97)
<b>Tumor grade</b>	
I	30 (81.08)
II + III	7 (18.92)
<b>TNM stage</b>	
I + II	32 (86.49)
III + IV	5 (13.51)
<b>pT stage</b>	
T1 + T2	35 (94.59)
T3 + T4	2 (5.41)
<b>pN stage</b>	
N0	35 (94.59)
N1	2 (5.41)
<b>pM stage</b>	
M0	37 (100)
M1	0 (0)

The pink module was mainly involved in the CDC6 association with ORC (origin complex), hepatitis C, and transcriptional activator activity, and it was observed to be negatively associated with age at initial diagnosis and new tumor events. Currently, sporadic evidence has shown a connection between the CDC6-related pathway and PRCC. CDC6 is a regulator of the initiation of DNA replication in human cells<sup>36,37</sup>. CDC6 may act as an ATPase switch that binds to Mcm2-7 in CDT1:CDC6:ORC.

Human CDC6 protein levels decrease early in the G<sub>1</sub> phase but remain unchanged throughout the cell cycle<sup>36,38</sup>. It has been reported that after cells enter the S phase, CDC6 is phosphorylated and excluded from the nucleus, which allows it to be easily ubiquitinated and degraded. Supplementation of CDC6 protein levels during G<sub>1</sub> appears to be regulated by the E2F transcription factors<sup>38</sup>. In addition, the aberrant expression of genes involved in the regulation of transcriptional activator activity pathways could also influence tumorigenesis.

The yellow module was shown to involve genes mainly related to embryonic organ development, epithelial cell proliferation, renal system development, extracellular matrix, transcriptional activator activity, neuroactive ligand–receptor interaction, breast cancer, GPCR ligand binding, class A/1 signaling (rhodopsin-like receptors), and G alpha (q) signaling and was negatively associated with new tumor events and positively associated with the number of years of smoking. The yellow module-related pathways have been frequently reported to be involved in the tumor progression process. For example, the G protein coupled receptor (GPCR) superfamily is at the center of many different signaling pathways involved in various aspects of human physiology, and they are composed of approximately 800 different members<sup>39</sup>. Although they share a common seven-transmembrane (7TM) architecture<sup>40</sup>, they can recognize a variety of different ligands, including small molecules, peptides, lipids, and proteins, and may serve as potential therapeutic targets for cancer patients<sup>41–43</sup>.

Although the light yellow module could serve as a predictor for OS and RFS, the genes in this module may not be significantly enriched. Interestingly, the GO-Reactome results showed that genes in the light yellow module may be enriched in the inactivation, recovery, and regulation of the phototransduction cascade and the mine ligand-binding receptor pathways, and further functional studies are warranted to determine whether these pathways



**Figure 9.** Immunohistochemistry analysis for the PCDH12 expression in PRCC. A value of  $p < 0.05$  was regarded as statistically significant. \*\*\* $p < 0.001$ .

could influence PRCC initiation and progression. The salmon module was mainly enriched in the olfactory signaling and olfactory transduction pathways. We could only weakly connect the olfactory signaling or olfactory transduction pathways with cancers. Currently, emerging evidence suggests that the regulation of cell–cell recognition, exocytosis, migration, proliferation, apoptotic cycles, and pathfinding involve olfactory receptors and the related pathways<sup>44,45</sup>. In addition, there is growing evidence that olfactory receptors have great potential as diagnostic and therapeutic tools<sup>46</sup>. We also validated the stability of the hub genes using an external GEO dataset and found that most of the genes we identified were consistent with our findings. Furthermore, we employed the IHC assay and identified that the expression of PCDH12 was upregulated in PRCC patients with high pathological grade (II + III vs. I). However, until now, few studies have reported the function of PCDH12 in cancers. Future efforts are warranted to explore it.

In summary, using WGCNA and other methods combined with clinical data from PRCC patients to analyze the RNA-seq data, a set of biomarkers that can predict the prognosis of PRCC patients were identified. These results have important clinical implications that will contribute to personalized treatment.

**ACKNOWLEDGMENTS:** *The presented research was financially supported by the National Science Foundation for Young Scientists (81400757 and 81802827), the National Natural Science Foundation of China (31430028, 81630019, and 81870519), the Scientific Research Foundation of the Institute for Translational Medicine of Anhui Province (2017ZHXY02), and the Natural Science Foundation of Guangdong Province, China (2017A030313800). X.F., M.Z., Y.W., Y.L., and S.F. designed the experiment. X.F., M.Z., J.M., and C.L. performed the research and analyzed the data. X.F. and M.Z. wrote the manuscript, and then S.F. and C.L. revised the manuscript. All authors read and approved the final manuscript. The authors declare no conflicts of interest.*

## REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA-Cancer J Clin.* 2019;69(1).
2. Antonio LB, Marina S, Rodolfo M, Ziya K. 2004 WHO classification of the renal tumors of the adults. *Eur Urol.* 2006;49(5):798–805.
3. Kovacs G, Akhtar M, Beckwith BJ, Bugert P, Cooper CS, Delahunt B, Eble JN, Fleming S, Ljungberg B, Medeiros LJ, Moch H, Reuter VE, Ritz E, Roos G, Schmidt D, Srigley JR, Störkel S, van den Berg E, Zbar B. The Heidelberg classification of renal cell tumours. *J Pathol.* 1997;183(2):131–3.
4. Egbert O, W Kimryn R, Kerstin J, A Rose B, Frédéric P, Finley DS, Mulders PFA, Ziya K, Hirotsugo U, Arie B. Basic research in kidney cancer. *Eur Urol.* 2011;60(4):622–33.
5. Rini BI, Campbell SC, Escudier B. Renal cell carcinoma. *Lancet* 2009;373(9669):1119–32.
6. Störkel S, Eble JN, Adlaka K, Amin M, Blute ML, Bostwick DG, Darson M, Delahunt B, Iczkowski K. Classification of renal cell carcinoma: Workgroup No. 1. Union Internationale Contre le Cancer (UICC) and the American Joint Committee on Cancer (AJCC). *Cancer* 1997;80(5):987–9.
7. Akhtar M, Al-Bozom IA, Al Hussain T. Papillary renal cell carcinoma (PRCC): An update. *Adv Anat Pathol.* 2019;26(2):124–32.
8. Vrdoljak E, Ciuleanu T, Kharkevich G, Mardiak J, Mego M, Padrik P, Petruželka L, Purkalne G, Shparyk Y, Skrbinc B, Szczylik C, Torday L. Optimizing treatment for patients with metastatic renal cell carcinoma in the Central and Eastern European region. *Expert Opin Pharmacother.* 2012;13(2):159–74.
9. Chan JY, Choudhury Y, Tan MH. Predictive molecular biomarkers to guide clinical decision making in kidney cancer: Current progress and future challenges. *Expert Rev Mol Diagn.* 2015;15(5):631–46.
10. Jacobsen J, Grankvist KT, Bergh A, Landberg G, Ljungberg B. Expression of vascular endothelial growth factor protein in human renal cell carcinoma. *BJU Int.* 2004;93(3):297–302.
11. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009;458(7239):719–24.
12. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp Oncol. (Pozn)* 2015;19(1A):A68–77.
13. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* 2008;9:559.
14. Plaisier CL, Horvath S, Huertas-Vazquez A, Cruz-Bautista I, Herrera MF, Tusie-Luna T, Aguilar-Salinas C, Pajukanta P. A systems genetics approach implicates USF1, FADS3, and other causal candidate genes for familial combined hyperlipidemia. *PLoS Genet.* 2009;5(9):e1000642.
15. Miller JA, Steve H, Geschwind DH. Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. *Proc Natl Acad Sci USA* 2010;107(28):12698–703.
16. Farber CR. Identification of a gene module associated with BMD through the integration of network analysis and genome-wide association data. *J Bone Miner Res* 2010;25(11):2359–67.
17. Chen P, Wang F, Feng J, Zhou R, Chang Y, Liu J, Zhao Q. Co-expression network analysis identified six hub genes in association with metastasis risk and prognosis in hepatocellular carcinoma. *Oncotarget* 2017;8(30):48948–58.
18. He Z, Sun M, Ke Y, Lin R, Xiao Y, Zhou S, Zhao H, Wang Y, Zhou F, Zhou Y. Identifying biomarkers of papillary renal cell carcinoma associated with pathological stage by weighted gene co-expression network analysis. *Oncotarget* 2017;8(17):27904–14.
19. Justo N, Wilking N, Jonsson B, Luciani S, Cazap E. A review of breast cancer care and outcomes in Latin America. *Oncologist* 2013;18(3):248–56.
20. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
21. Fuller T, Langfelder P, Presson A., Horvath S. Review of weighted gene coexpression network analysis. In: Lu HS, Schölkopf B., Zhao H, editors. *Handbook of statistical bioinformatics*. Berlin, Germany: Springer Handbooks of Computational Statistics; 2011.
22. Ghazalpour A, Doss S, Zhang B, Wang S, Plaisier C, Castellanos R, Brozell A, Schadt EE, Drake TA, Lusk AJ, Horvath S. Integrating genetic and network analysis to

- characterize genes related to mouse weight. *PLoS Genet.* 2006;2(8):e130.
23. Alter O, Brown PO, Botstein D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci USA* 2000;97(18):10101–6.
  24. Langfelder P, Horvath S. Eigengene networks for studying the relationships between co-expression modules. *BMC Syst Biol.* 2007;1:54.
  25. Lin H, Zelterman D. Modeling survival data: Extending the Cox model. *Technometrics* 2000;44(1):85–6.
  26. Kassambara A. Drawing survival curves using ‘ggplot2’ [R package survminer version 0.2.0]. 2017.
  27. Zhang L, Liu Y, Chen XG, Zhang Y, Chen J, Hao ZY, Fan S, Zhang LG, Du HX, Liang CZ. MicroRNA expression profile in chronic nonbacterial prostatitis revealed by next-generation small RNA sequencing. *Asian J Androl.* 2019;21(4):351–9.
  28. Pflaum T, Hausler T, Baumung C, Ackermann S, Kuballa T, Rehm J, Lachenmeier DW. Carcinogenic compounds in alcoholic beverages: An update. *Arch Toxicol.* 2016;90(10):2349–67.
  29. Xu M, Luo J. Alcohol and cancer stem cells. *Cancers (Basel)* 2017;9(11).
  30. Gu S, Nguyen BN, Rao S, Li S, Shetty K, Rashid A, Shukla V, Deng CX, Mishra L, Mishra B. Alcohol, stem cells and cancer. *Genes Cancer* 2017;8(9–10):695–700.
  31. Machida K, Chen CL, Liu JC, Kashiwabara C, Feldman D, French SW, Sher L, Hyeongnam JJ, Tsukamoto H. Cancer stem cells generated by alcohol, diabetes, and hepatitis C virus. *J Gastroenterol Hepatol.* 2012;27(Suppl 2): 19–22.
  32. Oberst A, Dillon CP, Weinlich R, McCormick LL, Fitzgerald P, Pop C, Hakem R, Salvesen GS, Green DR. Catalytic activity of the caspase-8–FLIP(L) complex inhibits RIPK3-dependent necrosis. *Nature* 2011;471(7338):363–7.
  33. Kaiser WJ, Upton JW, Long AB, Livingston-Rosanoff D, Daley-Bauer LP, Hakem R, Caspary T, Mocarski ES. RIP3 mediates the embryonic lethality of caspase-8-deficient mice. *Nature* 2011;471(7338):368–72.
  34. Su Z, Yang Z, Xie L, DeWitt JP, Chen Y. Cancer therapy in the necroptosis era. *Cell Death Differ.* 2016;23(5):748–56.
  35. Galluzzi L, Kepp O, Chan FK, Kroemer G. Necroptosis: Mechanisms and relevance to disease. *Annu Rev Pathol.* 2017;12:103–30.
  36. Drury LS, Perkins G, Diffley JF. The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. *Curr Biol.* 2000;10(5):231–40.
  37. Saha P, Chen J, Thome KC, Lawlis SJ, Hou ZH, Hendricks M, Parvin JD, Dutta A. Human CDC6/Cdc18 associates with Orc1 and cyclin-cdk and is selectively eliminated from the nucleus at the onset of S phase. *Mol Cell Biol.* 1998;18(5):2758–67.
  38. Yan Z, DeGregori J, Shohet R, Leone G, Stillman B, Nevins JR, Williams RS. Cdc6 is regulated by E2F and is essential for DNA replication in mammalian cells. *Proc Natl Acad Sci USA* 1998;95(7):3603–8.
  39. Hauser AS, Attwood MM, Rask-Andersen M, Schioth HB, Gloriam DE. Trends in GPCR drug discovery: New agents, targets and indications. *Nat Rev Drug Discov.* 2017;16(12):829–42.
  40. Bockaert J, Pin JP. Molecular tinkering of G protein-coupled receptors: An evolutionary success. *EMBO J.* 1999;18(7):1723–9.
  41. Barton M, Filardo EJ, Lolait SJ, Thomas P, Maggiolini M, Prossnitz ER. Twenty years of the G protein-coupled estrogen receptor GPER: Historical and personal perspectives. *J Steroid Biochem Mol Biol.* 2018;176:4–15.
  42. Guidolin D, Marcoli M, Tortorella C, Maura G, Agnati LF. Receptor–receptor interactions as a widespread phenomenon: Novel targets for drug development? *Front Endocrinol. (Lausanne)* 2019;10:53.
  43. Ma X, Xiong Y, Lee LTO. Application of nanoparticles for targeting G protein-coupled receptors. *Int J Mol Sci.* 2018;19(7). pii: E2006.
  44. Durrant DM, Ghosh S, Klein RS. The olfactory bulb: An immunosensory effector organ during neurotropic viral infections. *ACS Chem Neurosci.* 2016;7(4):464–9.
  45. Maßberg D, Hatt H. Human olfactory receptors: Novel cellular functions outside of the nose. *Physiol Rev.* 2018;98(3):1739–63.
  46. Kleene SJ. The electrochemical basis of odor transduction in vertebrate olfactory cilia. *Chem Senses* 2008;33(9):839–59.