

# Construction and Validation of a Robust Genomic Classifier for High-Risk Sporadic Type 2 Papillary Renal Cell Carcinoma: Support More Accurate Utilization of mTOR inhibitors

Xi Tian (✉ [Tianxistudying@163.com](mailto:Tianxistudying@163.com))

Fudan University Shanghai Cancer Center

**Wen-Hao Xu**

Fudan University Shanghai Cancer Center

**Aihetaimujiang Anwaier**

Department of Urology, Fudan University Shanghai Cancer Center, Shanghai

**Yue Wang**

Fudan University Shanghai Cancer Center

**Hua-Lei Gan**

Fudan University Shanghai Cancer Center

**Wen-Jie Luo**

Fudan University Shanghai Cancer Center

**Hong-Kai Wang**

Fudan University Shanghai Cancer Center

**Fang-Ning Wan**

Fudan University Shanghai Cancer Center

**Yu Zhu**

Fudan University Shanghai Cancer Center

**Da-Long Cao**

Fudan University Shanghai Cancer Center

**Guo-Hai Shi**

Fudan University Shanghai Cancer Center

**Yuan-Yuan Qu**

Fudan University Shanghai Cancer Center

**Hai-Liang Zhang**

Fudan University Shanghai Cancer Center

**Ding-Wei Ye**

Fudan University Shanghai Cancer Center

## Research

**Keywords:** sporadic type 2 papillary renal cell carcinoma, genomic classifier, prognosis, everolimus

**Posted Date:** March 4th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-264118/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

To construct a robust genomic classifier for high-risk sporadic type 2 papillary renal cell carcinoma (sPRCC2) and identify potential therapeutic targets. A cohort from The Cancer Genome Atlas and two datasets from Gene Expression Omnibus were examined. Common differentially expressed genes were screened, and the ConsensusClusterPlus package was used to identify potential high-risk molecular subtypes of sPRCC2. Targeting protein for Xklp2 (TPX2) expression was used to simulate the classifier according to the logistic model. Ninety-two samples from Fudan University Shanghai Cancer Center (FUSCC) were obtained, and TPX2 immunostaining was performed to validate the predictive ability of the genomic classifier. Retrospective analysis was used to compare drug efficacy between the groups. High-risk sPRCC2 had worse overall (hazard ratio [HR] = 7.804) and disease-free survival (HR = 8.777) than low-risk sPRCC2, and the tumor and lymph node stages were significantly higher in high-risk sPRCC2. Gene set enrichment analysis revealed significant enrichment of genes in the mammalian target of rapamycin (mTOR) complex 1 signaling pathway in the high-risk group. In the FUSCC cohort, high-risk sPRCC2 (high TPX2 expression) was significantly correlated with worse overall and progression-free survival. Retrospective analysis indicated that mTOR inhibitor (everolimus) had greater efficacy in the high-risk group than in the low-risk group (overall response rate: 28.6% vs. 16.7%) and that everolimus had greater efficacy than sunitinib in the high-risk group (overall response rate: 28.6% vs. 20%). This study successfully constructed a genomic classifier for identifying high-risk sPRCC2. mTOR inhibitors may have good efficacy in patients with high-risk sPRCC2.

## Introduction

Renal cell carcinoma (RCC) is the third most common malignant tumor of the genitourinary system. In 2019, 73,750 people were newly diagnosed with renal tumors, and 14,830 deaths were attributed to renal tumors[1]. Clear cell RCC (ccRCC) represents approximately 70% of kidney cancer cases in adults[2]. Papillary renal cell carcinoma (PRCC) is the most common non-clear cell RCC (nccRCC), accounting for 10–15% of RCCs[3]. Delahunt and Eble[4] characterized the histologic dissimilarities of PRCC and divided this malignancy into two subtypes (PRCC1 and PRCC2). Molecular analysis further clarified differences between the two subtypes. PRCC1 features gains in chromosomes 7, 17, 16 and 20 but loss of the Y chromosome[5]. MET pathway activation is frequently implicated in PRCC1[6]. Conversely, PRCC2 has a more heterogenous spectrum of chromosomal gains and losses. It has been reported that 8q gains are especially related to the poor prognosis of PRCC2, and the NRF–ARE2 pathway was also revealed to be enriched in PRCC2[7, 8]. Previous studies demonstrated that PRCC2 has significantly worse clinical outcomes than PRCC1[9, 10]. In summary, PRCC2 differs from PRCC1 and features a more aggressive phenotype.

PRCC2 can be further divided into hereditary and sporadic types. The hereditary form is associated with biallelic inactivation of the gene encoding the Krebs cycle enzyme fumarate hydratase (FH), which leads to hereditary leiomyomatosis and RCC (HLRCC) syndrome, which is characterized by a high incidence of RCC, uterine leiomyoma, and cutaneous leiomyomatosis [11, 12]. Patients with HLRCC syndrome are also

genetically susceptible to bladder cancer, collecting duct tumors, and adult Leydig cell tumors of the testes[13–15]. Sporadic PRCC2 (sPRCC2) accounts for most cases of PRCC2, and previous studies demonstrated that despite differences in genetic etiology, sPRCC2 shares many clinical and morphologic phenotypes with HLRCC syndrome[16]. The most prominent common biochemical feature of HLRCC syndrome and sPRCC2 is the continuous activation of NRF2, which is caused by intracellular fumaric acid accumulation attributable to fumarate hydratase (FH) inactivation[16], but the mechanism of NRF2 activation in sPRCC2 has not been determined.

Although rapid progress in medical science has facilitated the development of cancer therapy and multiple new drugs exert antitumor effects against ccRCC, problems remain in the management of PRCC2. Numerous clinical trials aimed to explore potential useful treatments for PRCC. Ravaud et al.[17] found that sunitinib was effective in the treatment of metastatic PRCC1 and PRCC2, but its efficacy was lower than that against metastatic ccRCC. Both progression-free survival (PFS) and overall survival (OS) are longer in PRCC1. In addition, Armstrong et al.[18] claimed that compared with everolimus, sunitinib improved PFS in patients with metastatic nccRCC. However, the results of these clinical trials including various targeted therapies and immunotherapies did not revolutionize the treatment of PRCC[19–22]. Because of its rarity and heterogeneity, there is little useful information regarding the rational clinical management of metastatic PRCC2.

Although OS is considered short in PRCC2, we also identified a subset of patients with histopathologically confirmed sPRCC2 and prolonged survival. It is important to clarify the mechanism responsible for the difference in survival. In this research, we focused on the molecular pattern of sPRCC2 and explored outcomes in different sPRCC2 subtypes using bioinformatics. A genomic classifier was successfully constructed, and the high-risk sPRCC2 subgroup was identified. Ninety-two samples from Fudan University Shanghai Cancer Center (FUSCC) were examined to validate the predictive ability of the genomic classifier, and retrospective analysis was used to compare drug efficacy between the groups.

## **Materials And Methods**

### **Comparison of PRCC1 and PRCC2 in The Cancer Genome Atlas (TCGA) cohort**

Data for 77 patients with PRCC1 and 85 patients with PRCC2 and complete genetic alteration and clinical data were obtained from TCGA. Clinical information and genetic alterations in PRCC1 and PRCC2 were obtained from cBioPortal (<https://www.cbioportal.org/>). The Kaplan–Meier method was used to compare OS and disease-free survival (DFS) between the PRCC1 and PRCC2 groups. The chi-squared test and Kruskal–Wallis test were also applied to assess other clinical information including American Joint Committee on Cancer tumor stage, lymph node stage, metastasis stage, and serum calcium levels.

### **Gene expression profiles of PRCC2 and differential gene expression analysis**

Gene expression profiles and clinical information for patients with PRCC in TCGA were downloaded from <https://portal.gdc.cancer.gov/>. Germline mutation data in PRCC2 were obtained from the supplementary file of a previous study[23]. Because the molecular patterns and clinical behavior varied between patients with hereditary PRCC2 (patients with *FH* germline mutation) and sPRCC2, this study only focused on sPRCC2 (82 samples from TCGA, Table 1) and hereditary PRCC2 was excluded (two patients). The mutation patterns and corresponding gene expression patterns of these 82 sPRCC2 samples were obtained from cBioPortal. Two datasets containing expression profiles of sPRCC2 were downloaded from Gene Expression Omnibus: GSE26574 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE26574>, contains 12 sPRCC2 samples) and GSE48352 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE48352>, contains 19 sPRCC2 samples). The limma package [24] and GEO2R were used to explore differentially expressed genes (DEGs) between normal and sPRCC2 tissues from these three cohorts (adjusted  $p < 0.05$  and fold change  $\geq 2$ ). A Venn diagram was applied to identify the overlapping upregulated and downregulated DEGs.

Table 1  
Clinicopathological characteristics of 82 patients diagnosed with sporadic type-2 papillary renal cell carcinoma (without FH germline mutation) from TCGA cohort

Characteristics	Entire cohort (N = 82)
N (%)	
Age	
< 70 years	51 (62.2)
≥ 70 years	30 (36.6)
Censored	1 (1.2)
Gender	
Male	58 (70.7)
Female	24 (29.3)
Laterality	
Left	46 (56.1)
Right	36 (43.9)
pathologic T stage†	
T1 – T2	58 (70.7)
T3 – T4	24 (29.3)
pathologic N stage†	
N0	15 (18.3)
N1	11 (13.4)
N2	3 (3.7)
NX & Censored	53 (64.6)
pathologic M stage†	
M0	31 (37.8)
M1	2 (2.4)
MX & Censored	49 (59.8)
† TNM scoring system: Tumor size, Lymph Nodes affected, Metastases.	

# Identifying the high-risk subtype of sPRCC2 and establishing the genomic classifier

Protein–protein interaction (PPI) networks of the overlapping upregulated and downregulated DEGs were separately constructed using the Search Tool for the Retrieval of Interacting Genes (<http://string-db.org>, version 10.0) online database. MCODE (version 1.4.2)[25], a Cytoscape plug-in[26], was used to identify the most significant hub genes in the PPI network. According to the gene expression of hub genes, the ConsensusClusterPlus[27] package was applied to identify potential molecular subtypes of sPRCC2, and a cumulative distribution function (CDF) curve was used to determine the most accurate number of clusters (relative change in the area under the CDF curve  $\geq 0.1$  was considered significant). L-clusters and H-clusters were separately used to represent identified clusters according to the expression patterns of downregulated and upregulated hub genes. In this research, the numbers of L-clusters and H-clusters were set at 5 and 7, respectively, according to the CDF curves. The clusters were separately simplified to two main clusters according to the cluster dendrogram. The Kaplan–Meier method was used to compare OS and DFS between the clusters. The H-clusters were successfully simplified into two subtypes, and the subtype with significantly worse prognosis was defined as the high-risk group. Receiver operating characteristic[28] (ROC) curves were constructed to describe the binary classifier value of the classifier using the area under the curve (AUC).

## Comparison of low-risk and high-risk sPRCC2

The H-clusters were simplified into low-risk ( $n = 69$ ) and high-risk groups ( $n = 13$ ). The chi-squared test and Kruskal–Wallis test were also applied to compare clinical information and genomic variation between the two groups using cBioPortal. To explore the mechanism responsible for the worse prognosis of high-risk sPRCC2, we performed gene set enrichment analysis (GSEA) to identify potential differences in transcriptomics.

## Validating the genomic classifier using the FUSCC cohort

Because a genomic classifier that could identify high-risk sPRCC2 was constructed using the expression information of 14 upregulated hub genes, we aimed to simplify the genomic classifier. Bivariate logistic regression analysis was used, and we found that clusters separated using targeting protein for Xklp2 (*TPX2*) expression (cutoff was set as the thirteenth expression value) could fully match the clusters as divided by the genomic classifier. This study included 92 patients (clinical information is listed in Table 2) with histopathologically confirmed sPRCC2 (positive staining for FH) who underwent surgical treatment at FUSCC between 2009 and 2019, and tumor specimens were obtained with informed consent. Immunostaining of TPX2 was performed using a rabbit monoclonal anti-TPX2 antibody (Cat. ab270612, Abcam, USA). Positive or negative staining for a certain protein on a formalin-fixed, paraffin-embedded slide was independently assessed by two experienced pathologists. The staining intensity level was graded as follows: 0, no staining; 1, weak staining; 2, moderate staining; and 3, strong staining. The extent of staining ranged 0–4 based on the percentage of immunoreactive tumor cells (0%, 1–25%, 26–50%, 51–75%, 76–100%). The overall immunohistochemistry (IHC) score was obtained by multiplying the

staining intensity by the extent of staining. IHC scores of 0–3 represented low risk, and scores of 4–12 indicated high risk. Then, the Kaplan–Meier method was used to compare OS and PFS between the groups.

Table 2

Clinicopathological characteristics of 92 patients diagnosed with sporadic type-2 papillary renal cell carcinoma (FH-IHC: positive) from FUSCC cohort

<b>Characteristics</b>	<b>Entire cohort (N = 92)</b>
N (%)	
Age	
< 70 years	81(88.0)
≥ 70 years	11 (12.0)
Gender	
Male	61 (66.3)
Female	31 (33.7)
Laterality	
Left	52 (56.5)
Right	40 (43.5)
Tumor size	
< 4 cm	34 (37.0)
≥ 4 cm	58 (63.0)
pathologic T stage†	
T1 – T2	68 (73.9)
T3 – T4	24 (26.1)
pathologic N stage†	
N0	57 (62.0)
N1	25 (27.2)
NX	10 (10.9)
pathologic M stage†	
M0	50 (54.3)
M1	24 (26.1)
MX	18 (19.6)

† TNM scoring system: Tumor size, Lymph Nodes affected, Metastases.

Characteristics	Entire cohort (N = 92)
Fuhrman nuclear grade	
I-II	27 (29.3)
III-IV	65 (70.7)
† TNM scoring system: Tumor size, Lymph Nodes affected, Metastases.	

## Validation of the potential utility of the genomic classifier for accurate drug selection

In total, 24 patients in the FUSCC cohort had a pathologically confirmed diagnosis of metastatic sPRCC2.

We retrospectively collected the baseline characteristics, treatment details, and clinical outcomes of these patients by reviewing their electronic medical records, and the details were verified by two investigators. In the low-risk group, six patients received everolimus as first-line therapy. In the high-risk group, seven patients were treated with everolimus as the first-line therapy, and five patients were treated with sunitinib as the first-line therapy. Radiologic assessment was performed according to RECIST 1.1 criteria[29] to classify the best response to treatment as complete response (CR), partial response (PR), stable disease (SD), or progressive disease (PD).

## Results

### PRCC2 was more aggressive than PRCC1

Both OS and DFS were shorter in patients with PRCC2 than in those with PRCC1 (both  $p < 0.05$ , Fig. 1A–B), and the chi-squared test indicated that PRCC2 was often correlated with a higher tumor stage and lymph node stage (Fig. 1C–D). The somatic mutation pattern between PRCC1 and PRCC2 was diverse (Fig. 1E). PRCC1 had higher frequencies of *KMT2C* and *PCLO* mutation, whereas the most characteristic somatic alteration in PRCC1 was *MET* mutation. However, *MET* mutation was only detected in two PRCC2 samples. Meanwhile, PRCC2 had higher frequencies of *CUL3*, *SETD2*, and *PBRM1* mutation. In summary, PRCC2 had a higher frequency of pathogenic mutations.

### Some common genes may play a key role in the malignant phenotype of sPRCC2

TCGA cohort included samples from two patients with PRCC2 as indicated by *FH* germline mutation, and the remaining 82 patients with PRCC2 were grouped into the sPRCC2 cohort (Fig. 2A). *CUL3* mutation was most common in patients with sPRCC2, and *SETD2*, *PBRM1*, and *KMT2C* mutations were also common. These somatic mutations inevitably exerted influences on the gene expression pattern of sPRCC2 (Fig. 2B). The GSE26574 and GSE48352 datasets were also used to explore DEGs between

sPRCC2 and normal tissues (Fig. 2C–D). In total, 316 downregulated and 65 upregulated genes were identified (Fig. 2E–F).

## **Genomic classifier could not identify the subtype with worse prognosis in L-clusters**

A PPI network was constructed, and downregulated hub genes (*BDKRB2*, *NPY1R*, *SUCNR1*, *KNG1*, *PTGER3*, *S1PR3*, *S1PR1*) were screened (Fig. 3A–B). L-clusters were divided into five clusters (L-clusters A–E), and the survival curve of each cluster was drawn (Fig. 3C–E). L-clusters were then simplified into two clusters according to the cluster dendrogram. The two subtypes did not exhibit significant differences in either OS or DFS (both  $p > 0.05$ ).

## **Genomic classifier identified a high-risk subtype of sPRCC2 in H-clusters**

A PPI network was constructed, and upregulated hub genes (*AURKA*, *TPX2*, *UBE2C*, *KIF4A*, *BUB1B*, *RRM2*, *CDC20*, *PTTG1*, *MELK*, *NUSAP1*, *TTK*, *CCNB2*, *CCNB1*, *TOP2A*) were identified (Fig. 4A–B). H-clusters were divided into seven clusters (H-clusters A–G), and the survival curve of each cluster was drawn (Fig. 4C–E). H-clusters were also simplified into two clusters according to the cluster dendrogram. One molecular subtype exhibited a significant worse prognosis (Fig. 4F–G) than the other subtype for both OS (HR = 7.804,  $p < 0.0001$ ) and DFS (HR = 8.777,  $p < 0.0001$ ). The ROC curve (AUC = 0.617) indicated that the classifier has favorable predictive ability for OS (Fig. 4H).

## **High-risk sPRCC2 is a highly aggressive molecular subtype**

The genomic classifier divided the TCGA cohort into two subtypes (Fig. 5A–C), namely high-risk ( $n = 69$ ) and low-risk groups ( $n = 13$ ). Chi-squared tests indicated that the high-risk group had a higher tumor stage, a higher lymph node stage, a higher frequency of new neoplasm events, lower hemoglobin levels, and a relative higher genomic alteration frequency (all  $p < 0.05$ , Fig. 5D–H). GSEA indicated that compared with the findings in the low-risk group, gene expression was significantly enriched for E2F targets, the G2M checkpoint, Myc targets, and other pathways in the high-risk group (Fig. 5I). To our interest, gene expression in the high-risk group was also enriched in the mammalian target of rapamycin (mTOR) complex 1 (mTORC1) signaling pathway (Fig. 5J), which may shed light on the accurate use of mTOR inhibitors such as everolimus.

## ***The genomic classifier also identified the subtype in the FUSCC cohort with the worse outcome and shed light on the accurate use of everolimus in sPRCC2***

As described in the methods, TPX2 expression level was used in place of the genomic classifier (Fig. 6A–B). Ninety-two patients with histopathologically confirmed sPRCC2 (positive staining for FH) were included (Fig. 6C–D), and images of TPX2 expression (IHC, low and high) are presented in Fig. 6E–F.

Samples with low and high TPX2 expression comprised the low-risk (N = 49) and high-risk groups (N = 43), respectively. Survival analysis indicated that both OS (HR = 3.361,  $p < 0.0001$ , Fig. 6G) and PFS (HR = 4.209,  $p < 0.0001$ , Fig. 6H) were significantly worse in the high-risk group than in the low-risk group. Although not statistically significant, among patients who received first-line everolimus therapy, PFS was better in the high-risk group (N = 7) than in the low-risk group (N = 6, Fig. 6I). A retrospective analysis also indicated that everolimus exhibited better efficacy (Fig. 6J) in the high-risk group than sunitinib (N = 5). In summary, everolimus displayed greater efficacy in the high-risk group than in the low-risk group (overall response rate: 28.6% vs. 16.7%), and everolimus had greater efficacy than sunitinib in the high-risk group, including a better overall response rate (28.6% vs. 20%) and greater reduction of the target lesion (Fig. 6K).

## Discussion

The present study successfully constructed a genomic classifier for identifying patients with high-risk sPRCC2 using transcriptomic data from multiple cohorts. The high-risk group had a significantly worse prognosis than the low-risk group concerning both OS and DFS. In addition, the high-risk group featured a higher tumor stage, a lymph node stage, a higher frequency of new neoplasm events, lower hemoglobin levels, and a relative higher genomic alteration frequency. The GSEA results indicated that compared with the findings in the low-risk group, gene expression in the high-risk group was significantly enriched in the mTORC1 signaling pathway, which may shed light on the use of mTOR inhibitors. In the external validation, TPX2 expression was used to simulate the classifier according to the logistic regression results, and it could fully match the clusters. This method for simplifying the classifier may be not strictly rigorous, but it also revealed its strong ability to predict OS and PFS in the FUSCC cohort. The stronger efficacy effect of everolimus in the high-risk sPRCC2 group, although not statistically significant, exceeded our expectation.

Previous studies demonstrated that PRCC2 represents a heterogeneous group of lesions that can be divided into various subtypes according to genetic and molecular patterns, and these patterns reflect differences in the clinical course and prognosis of the disease. In a previous study[30], comprehensive genomic profiling was performed to sequence 315 genes, and the commonly altered genes in PRCC2 were CDKN2A/B (18%), TERT (18%), NF2 (13%), and FH (13%). Yang et al.[31] identified two highly distinct molecular PRCC subclasses via morphologic correlation, and they found that G1-S and G2-M checkpoint genes were dysregulated in class 1 and class 2 tumors. A similar pattern was observed in this research. We found that gene expression in high-risk sPRCC2 was enriched in the G2M checkpoint pathway (Fig. 5I), which suggests that the G2M checkpoint plays a key role in the malignant phenotype of sPRCC2. In 2016, the TCGA research network[6] revealed that PRCC2 can be further classified into three individual subgroups based on molecular differences associated with patient survival. In this research, we compared the previously reported classifier with that developed in this study. The predictive accuracy of the two classifiers for OS was similar (**Figure S1**).

Our results demonstrated that high-risk sPRCC2 was significantly correlated with a higher tumor stage, a higher lymph node stage, and worse OS. Because the treatment of advanced PRCC2 remains difficult, it is of great importance to identify potential targets suppressing PRCC2 growth. As mentioned in a previous study[6], the classification of PRCC may have a significant impact on clinical and therapeutic management and clinical trial design. Mutation of NF2 (the Hippo pathway tumor suppressor) was observed in a number of PRCCs, and this pathway has been targeted in other cancers[32]. The NRF2–ARE pathway was upregulated in both hereditary PRCC and sPRCC2. Currently, researchers are interested in the NRF2–ARE pathway, and novel strategies targeting this pathway have recently been developed[33, 34]. In this study, we found that high-risk sPRCC2 exhibited excellent mTORC1 signaling pathway activity, which suggests the potential accurate use of mTOR inhibitors. Everolimus, an oral mammalian mTOR inhibitor, has antitumor activity in multiple cancer types[35], and previous research demonstrated that everolimus has some clinical benefit in patients with metastatic PRCC[36, 37]. In our retrospective analysis of the FUSCC cohort, everolimus exhibited a stronger drug effect against high-risk sPRCC2 than against low-risk sPRCC2, and everolimus had greater activity in the high-risk group than sunitinib. This result conferred that the genomic classifier can also guide the accurate use of everolimus in sPRCC2.

This study had several limitations. The nature of retrospective research limits the clinical value of this work. Further validation in multicenter or prospective studies is needed to verify the findings. However, it is difficult to conduct randomized controlled trials in sPRCC2 because of its rarity. There is also an urgent need for in vitro and in vivo experiments to explore the underlying mechanisms behind the genomic classifier.

## Abbreviations

RCC

renal cell carcinoma

ccRCC

clear cell renal cell carcinoma

nccRCC

non–clear cell renal cell carcinoma

PRCC

papillary renal cell carcinoma

PRCC1

type 1 papillary renal cell carcinoma

PRCC2

type 2 papillary renal cell carcinoma

sPRCC2

sporadic type 2 papillary renal cell carcinoma

TCGA

from the Cancer Genome Atlas

FUSCC

Fudan University Shanghai Cancer Center  
GEO  
Gene Expression Omnibus  
GSEA  
gene set enrichment analysis

## Declarations

**Acknowledgements:** We thank the TCGA databases and GEO (ID: GSE26574, GSE48352) for providing PRCC gene expression profiles.

**Consent for publication**☐Not applicable

**Author contributions:** The work presented here was carried out in collaboration among all authors. YDW, ZHL and QYY defined the theme of the study and discussed analysis, interpretation and presentation. TX, XWH and Aihetaimujiang drafted the manuscript, analyzed the data, developed the algorithm, and explained the results. WY, LWJ, WHK and WFN, participated in the collection of relevant data and helped draft the manuscript. ZY and CDL helped to perform the statistical analysis. SGH helped revise the manuscript and provided guiding suggestions. All the authors read and approved the final manuscript.

**Conflict of Interests:** The authors declare no competing interests.

**Ethics approval and consent to participate:** The Ethics approval and consent to participate of the current study was approved and consented by the ethics committee of Fudan University Shanghai Cancer center.

**Availability of data and material:** GSE26574 data sets was obtained from <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE26574>. GSE48352 data sets was obtained from <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE48352>. Gene expression profiles and clinical information of PRCC from TCGA were downloaded from <https://portal.gdc.cancer.gov/>. The data from FUSCC cohort during the current study available from the corresponding author on reasonable request.

**Funding:** This work is supported by Grants from the National Key Research and Development Project (No.2019YFC1316000) and National Natural Science Foundation of China (No.81772706 and No.81802525)

## References

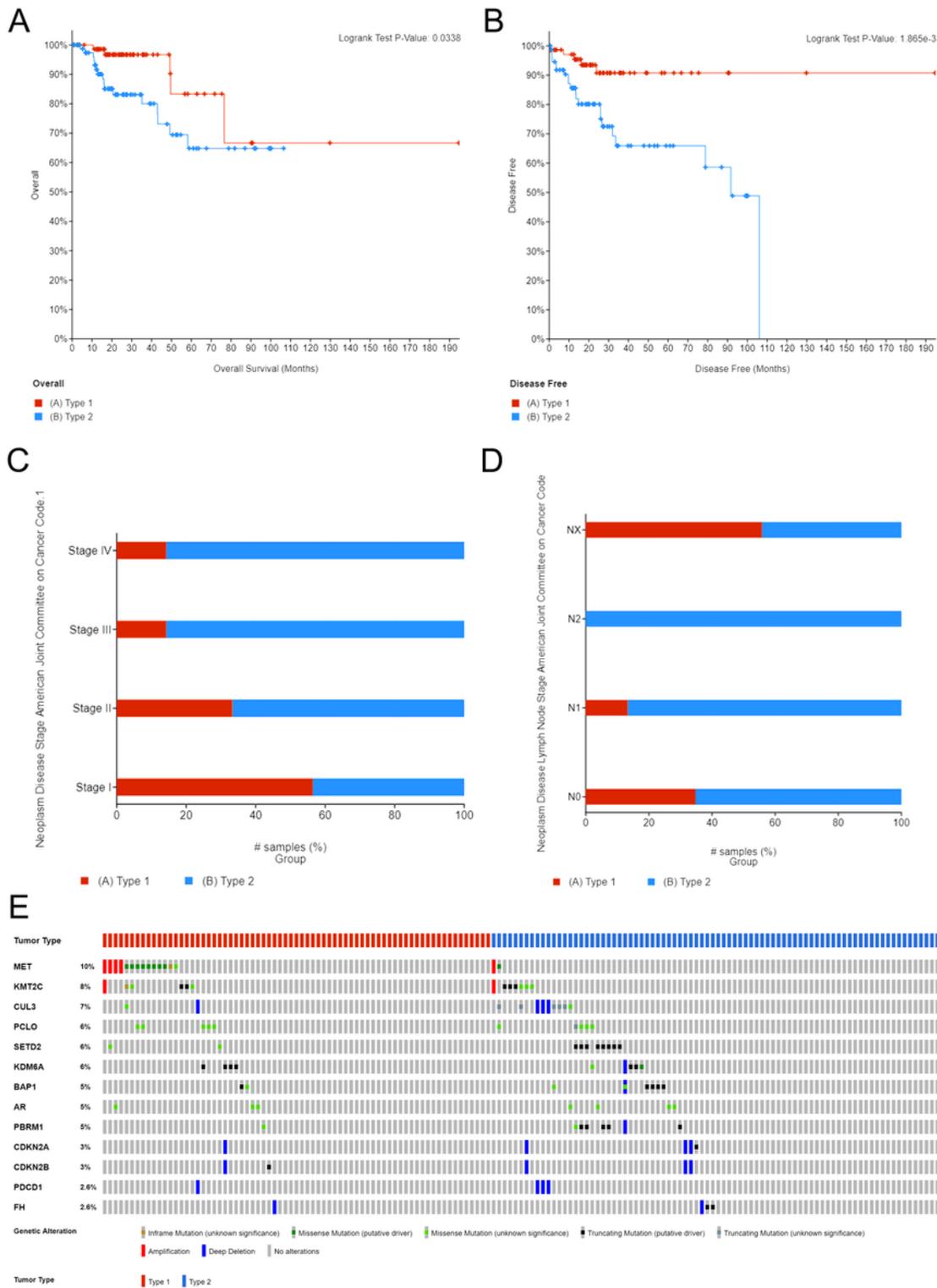
1. Siegel RL, Miller KD, Jemal A. *Cancer statistics, 2020*. CA: a cancer journal for clinicians, 2020. 70(1).
2. Gansler T, Fedewa S, Amin MB, et al., *Trends in reporting histological subtyping of renal cell carcinoma: association with cancer center type*. Human pathology, 2018. 74.

3. Choueiri TK, Vaishampayan U, Rosenberg JE, et al. Phase II and biomarker study of the dual MET/VEGFR2 inhibitor foretinib in patients with papillary renal cell carcinoma. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2013;31(2):181–6.
4. Delahunt B, Eble JN, McCredie MR, et al. Morphologic typing of papillary renal cell carcinoma: comparison of growth kinetics and patient survival in 66 cases. *Human pathology*. 2001;32(6):590–5.
5. Brunelli M, Eble JN, Zhang S, et al., *Gains of chromosomes 7, 17, 12, 16, and 20 and loss of Y occur early in the evolution of papillary renal cell neoplasia: a fluorescent in situ hybridization study*. *Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc*, 2003. **16**(10): p. 1053–1059.
6. Linehan WM, Spellman PT, Ricketts CJ, et al. Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma. *N Engl J Med*. 2016;374(2):135–45.
7. Furge KA, Chen J, Koeman J, et al. Detection of DNA copy number changes and oncogenic signaling abnormalities from gene expression data reveals MYC activation in high-grade papillary renal cell carcinoma. *Cancer research*. 2007;67(7):3171–6.
8. Marsaud A, Dadone B, Ambrosetti D, et al. Dismantling papillary renal cell carcinoma classification: The heterogeneity of genetic profiles suggests several independent diseases. *Genes Chromosom Cancer*. 2015;54(6):369–82.
9. Pignot G, Elie C, Conquy S, et al. Survival analysis of 130 patients with papillary renal cell carcinoma: prognostic utility of type 1 and type 2 subclassification. *Urology*. 2007;69(2):230–5.
10. Waldert M, Haitel A, Marberger M, et al. Comparison of type I and II papillary renal cell carcinoma (RCC) and clear cell RCC. *BJU Int*. 2008;102(10):1381–4.
11. Launonen V, Vierimaa O, Kiuru M, et al. Inherited susceptibility to uterine leiomyomas and renal cell cancer. *Proc Natl Acad Sci USA*. 2001;98(6):3387–92.
12. Tomlinson IPM, Alam NA, Rowan AJ, et al. Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nat Genet*. 2002;30(4):406–10.
13. Carvajal-Carmona LG, Alam NA, Pollard PJ, et al. Adult leydig cell tumors of the testis caused by germline fumarate hydratase mutations. *J Clin Endocrinol Metab*. 2006;91(8):3071–5.
14. Menko FH, Maher ER, Schmidt LS, et al. Hereditary leiomyomatosis and renal cell cancer (HLRCC): renal cancer risk, surveillance and treatment. *Familial cancer*. 2014;13(4):637–44.
15. Lehtonen HJ, Kiuru M, Ylisaukko-Oja SK, et al. Increased risk of cancer in patients with fumarate hydratase germline mutation. *J Med Genet*. 2006;43(6):523–6.
16. Ooi A, Wong J-C, Petillo D, et al. An antioxidant response phenotype shared between hereditary and sporadic type 2 papillary renal cell carcinoma. *Cancer cell*. 2011;20(4):511–23.
17. Ravaud A, Oudard S, De Fromont M, et al., *First-line treatment with sunitinib for type 1 and type 2 locally advanced or metastatic papillary renal cell carcinoma: a phase II study (SUPAP) by the French*

- Genitourinary Group (GETUG)†. Annals of oncology: official journal of the European Society for Medical Oncology, 2015. 26(6): p. 1123–1128.*
18. Armstrong AJ, Halabi S, Eisen T, et al. Everolimus versus sunitinib for patients with metastatic non-clear cell renal cell carcinoma (ASPEN): a multicentre, open-label, randomised phase 2 trial. *Lancet Oncol.* 2016;17(3):378–88.
  19. Irshad T, Olencki T, Zynger DL, et al. Bevacizumab in metastatic papillary renal cell carcinoma (PRCC). *J Clin Oncol.* 2011;29(15\suppl):e15158–8.
  20. Buti S, Bersanelli M, Maines F, et al., *First-Line Pazopanib in Non-clear-cell Renal carcinoma: The Italian Retrospective Multicenter PANORAMA Study.* *Clinical genitourinary cancer, 2017. 15(4): p. e609-e614.*
  21. Koshkin VS, Barata PC, Zhang T, et al. Clinical activity of nivolumab in patients with non-clear cell renal cell carcinoma. *J Immunother Cancer.* 2018;6(1):9.
  22. McKay RR, Bossé D, Xie W, et al. The Clinical Activity of PD-1/PD-L1 Inhibitors in Metastatic Non-Clear Cell Renal Cell Carcinoma. *Cancer immunology research.* 2018;6(7):758–65.
  23. Huang K-L, Mashl RJ, Wu Y, et al., *Pathogenic Germline Variants in 10,389 Adult Cancers.* *Cell, 2018. 173(2).*
  24. Ritchie ME, Belinda P, Di W, et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies.* 2015(7): p. 7.
  25. Bandettini WP, Kellman P, Mancini C, et al., *MultiContrast Delayed Enhancement (MCOE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study.* 2012. 14(1): p. 83–83.
  26. Smoot ME, Ruscheinski OK, Wang J, Ideker P-L. T, *Cytoscape 2.8: new features for data integration and network visualization.*
  27. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics.* 2010;26(12):1572–3.
  28. Heagerty PJ, Saha-Chaudhuri pbP. *survivalROC: Time-dependent ROC curve estimation from censored survival data.* 2013.
  29. Eisenhauer EA, Therasse P, Bogaerts J, et al., *New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1).* *European journal of cancer (Oxford, England: 1990), 2009. 45(2): p. 228–247.*
  30. Pal SK, Ali SM, Yakirevich E, et al. Characterization of Clinical Cases of Advanced Papillary Renal Cell Carcinoma via Comprehensive Genomic Profiling. *European urology.* 2018;73(1):71–8.
  31. Yang XJ, Tan M-H, Kim HL, et al. A molecular classification of papillary renal cell carcinoma. *Cancer research.* 2005;65(13):5628–37.
  32. Johnson R, Halder G. The two faces of Hippo: targeting the Hippo pathway for regenerative medicine and cancer treatment. *Nat Rev Drug Discov.* 2014;13(1):63–79.

33. Sporn MB, Liby KT. NRF2 and cancer: the good, the bad and the importance of context. *Nat Rev Cancer*. 2012;12(8):564–71.
34. Sourbier C, Ricketts CJ, Matsumoto S, et al. Targeting ABL1-mediated oxidative stress adaptation in fumarate hydratase-deficient cancer. *Cancer cell*. 2014;26(6):840–50.
35. Hortobagyi GN, Chen D, Piccart M, et al. Correlative Analysis of Genetic Alterations and Everolimus Benefit in Hormone Receptor-Positive, Human Epidermal Growth Factor Receptor 2-Negative Advanced Breast Cancer: Results From BOLERO-2. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2016;34(5):419–26.
36. Koh Y, Lim HY, Ahn JH, et al. Phase II trial of everolimus for the treatment of nonclear-cell renal cell carcinoma. *Annals of oncology: official journal of the European Society for Medical Oncology*. 2013;24(4):1026–31.
37. Blank CU, Bono P, Larkin JMG, et al. Safety and efficacy of everolimus in patients with non-clear cell renal cell carcinoma refractory to VEGF-targeted therapy: Subgroup analysis of REACT. *J Clin Oncol*. 2012;30(5\suppl):402–2.

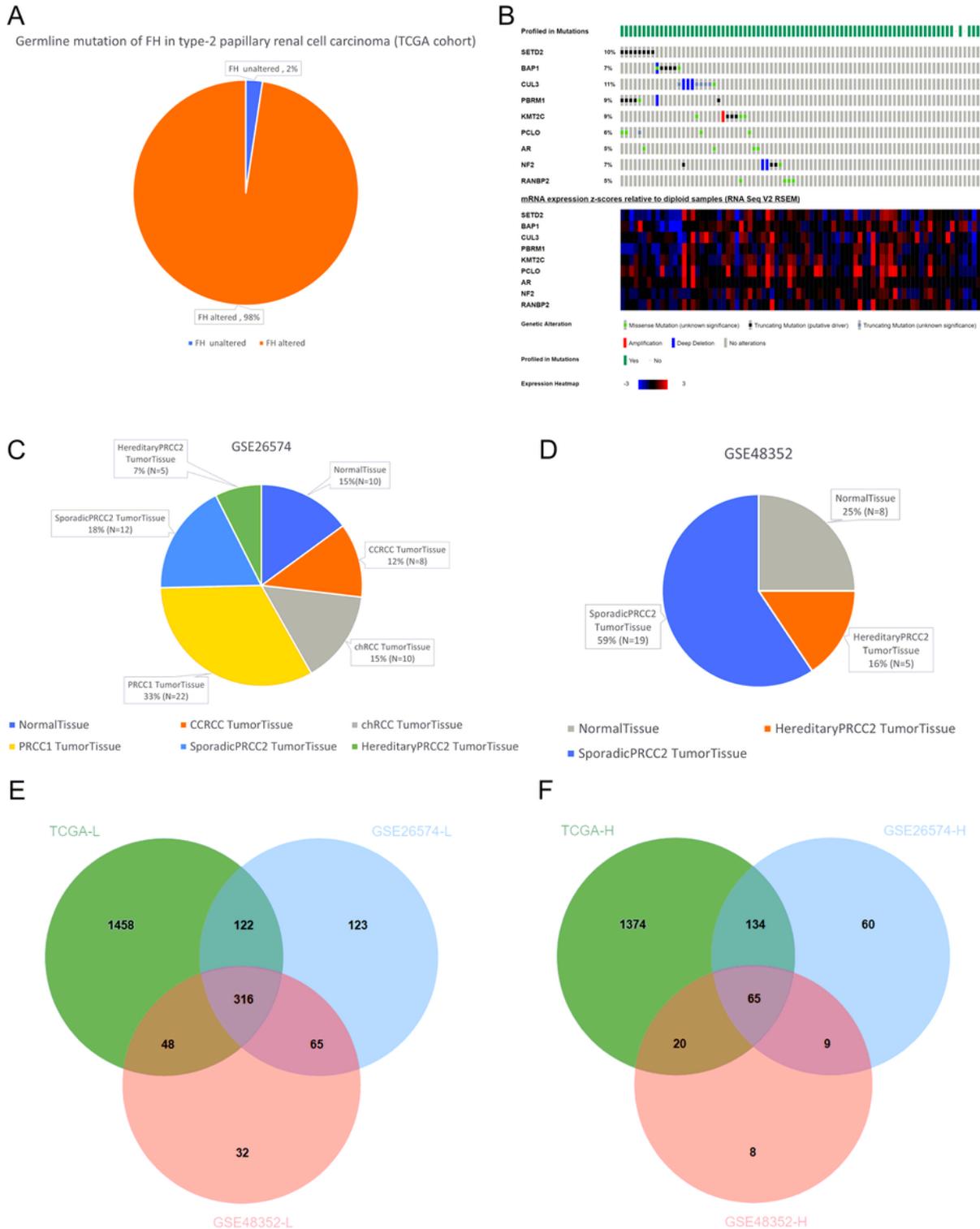
## Figures



**Figure 1**

Both the overall survival and disease free survival of PRCC2 was significantly shorter than PRCC1 ( $p < 0.05$ , A-B) and chi-squared test indicated that PRCC2 was often with higher tumor stage and lymph node stage (C-D). Somatic mutation pattern between PRCC1 and PRCC2 was diverse (E). PRCC1 has a higher frequency of KMT2C, PCLO and the most characteristic somatic alteration in PRCC1 was MET

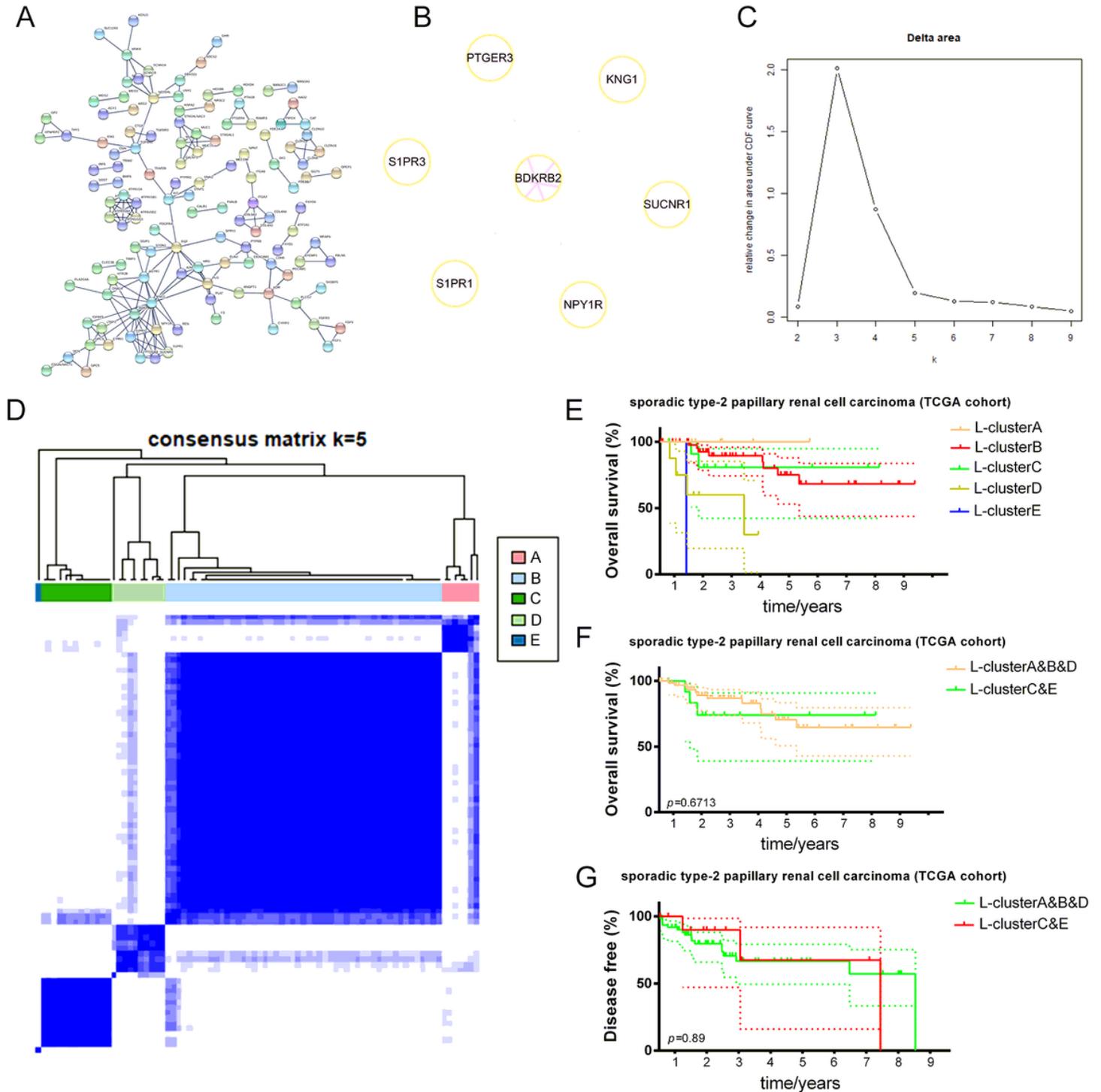
mutation. But there are just 2 samples with MET mutation in PRCC2. While PRCC2 has a higher frequency of CUL3, SETD2, PBRM1 mutation. In summary, PRCC2 has higher frequency of pathogenic mutations.



**Figure 2**

In patients enrolled from TCGA cohort, there are 2 PRCC2 patients with FH germline mutation and the other 82 PRCC2 patients were seen as sPRCC2 (A). CUL3 mutation has a highest frequency in sPRCC2 and SETD2, PBRM1, KMT2C also have a high mutation frequency. These somatic mutation inevitably

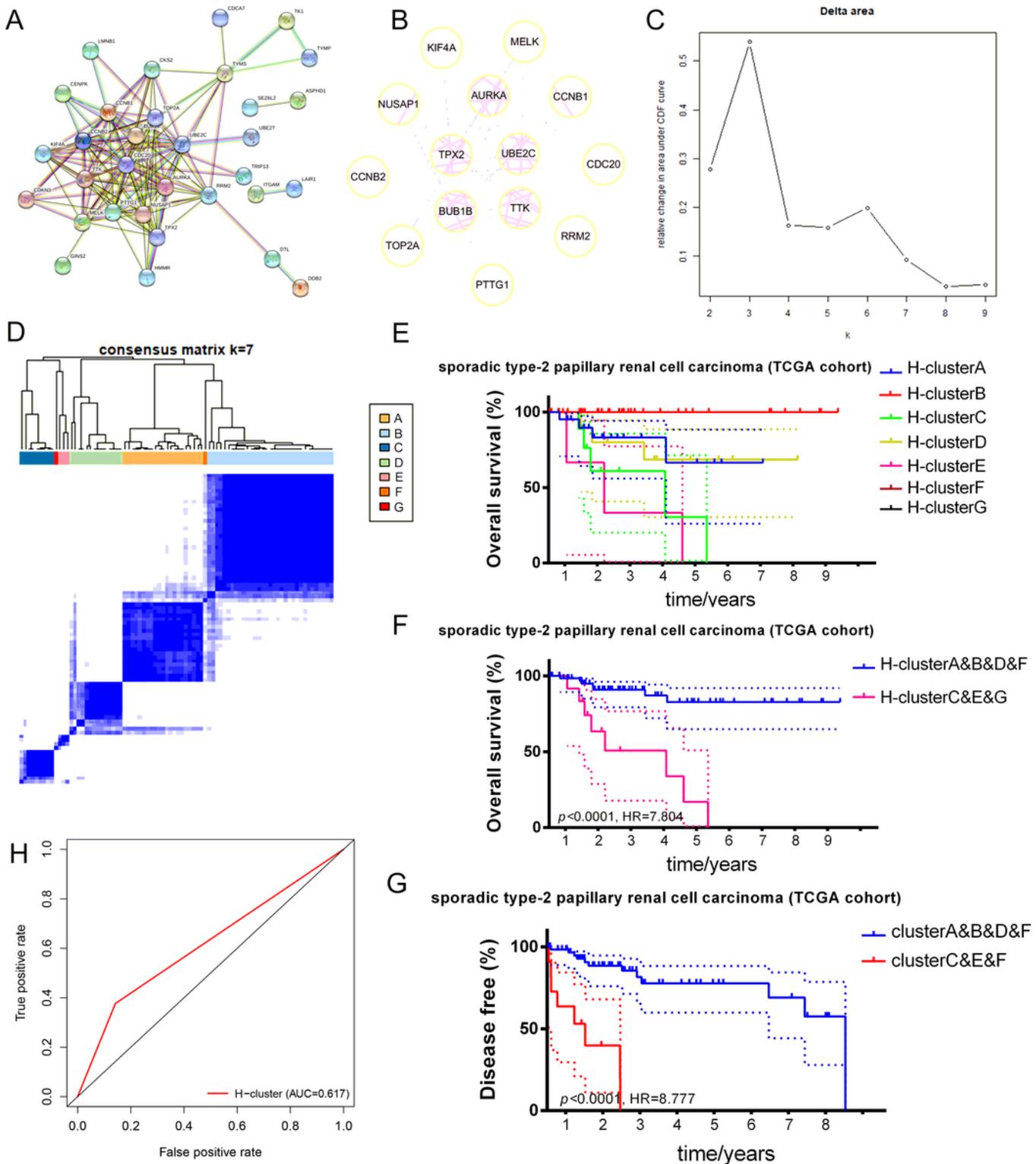
exerts influences on gene expression pattern of sPRCC2 (B). Apart from TCGA cohort, GSE26574 and GSE48352 were also enrolled to explore differentially expressed genes between sPRCC2 and normal tissues (C-D). 316 overlapped genes were identified as the common down-regulated genes and 65 overlapped genes were identified as the common up-regulated genes (E-F).



**Figure 3**

PPI network was constructed and down-regulated hub genes (BDKRB2, NPY1R, SUCNR1, KNG1, PTGER3, S1PR3, S1PR1) were screened (A-B). L-clusters were divided to five clusters (L-cluster A&B&C&D&E) and

survival curve of each cluster was drawn (C-E). L-clusters was then simplified to two cluster according to the cluster dendrogram. The two subtypes didn't show significant change in either overall survival or disease free survival (F-G,  $p > 0.05$ ).



**Figure 4**

PPI network was constructed and up-regulated hub genes (AURKA, TPX2, UBE2C, KIF4A, BUB1B, RRM2, CDC20, PTTG1, MELK, NUSAP1, TTK, CCNB2, CCNB1, TOP2A) were identified (A-B). H-clusters were

divided to seven clusters (H-cluster A-G) and survival curve of each cluster was drawn (C-E). H-clusters was also simplified to two cluster according to the cluster dendrogram. One of the molecular subtype showed significant worse prognosis (F-G) than the other subtype in both overall survival ( $p < 0.0001$ , HR=7.804) and disease free survival ( $p < 0.0001$ , HR=8.777). ROC curves (AUC=0.617) indicated that the classifier has a favorable predictive ability for overall survival (H).

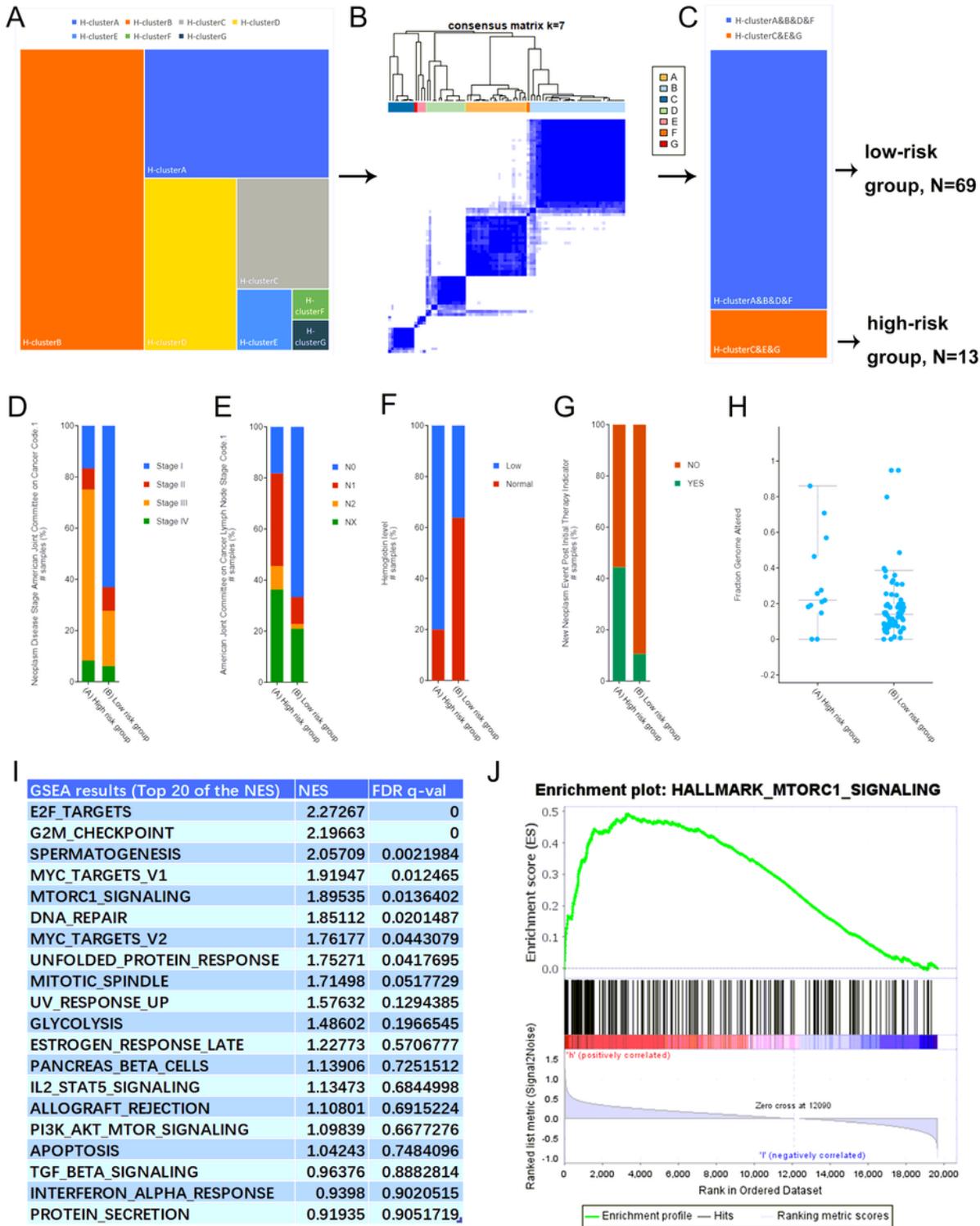
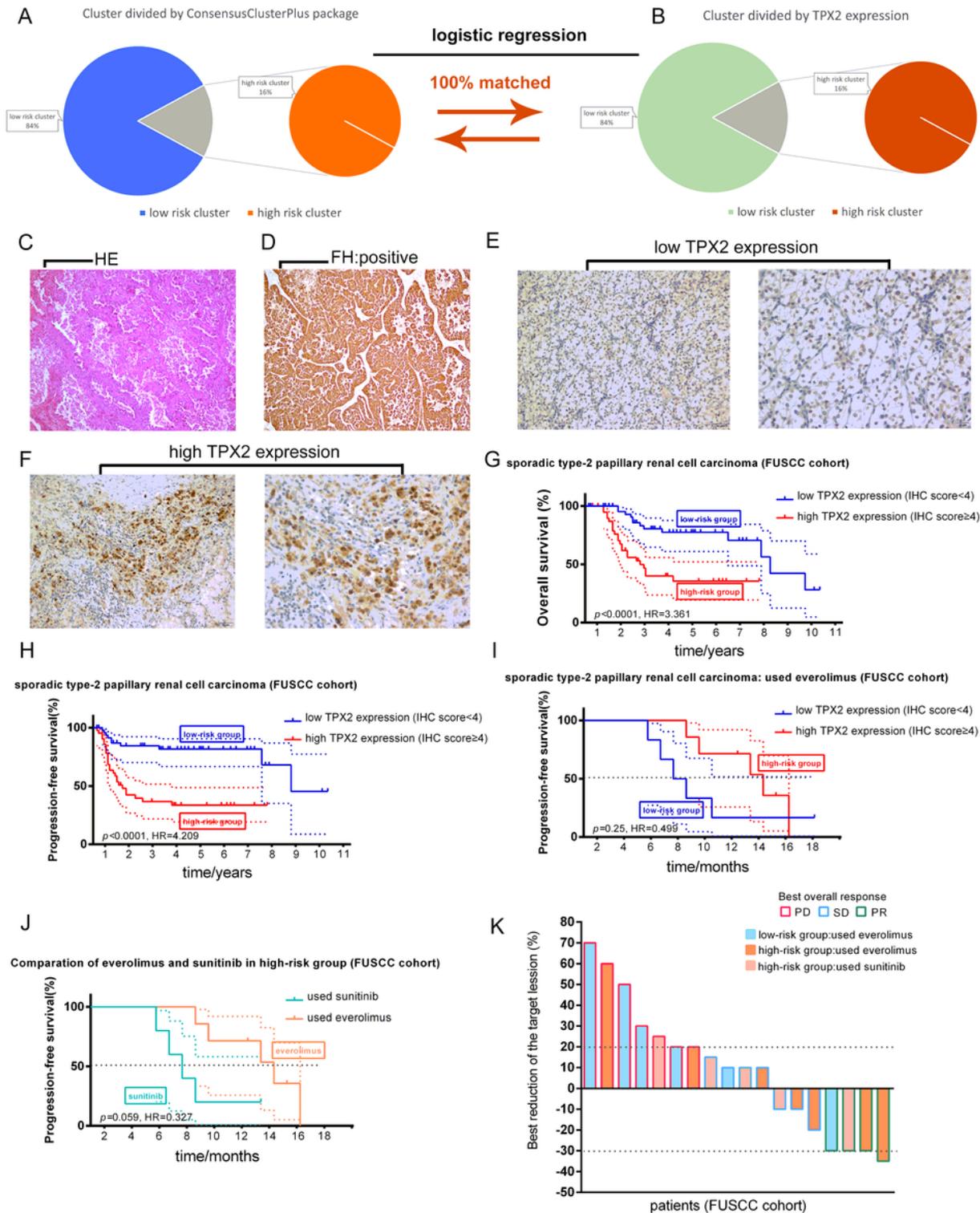


Figure 5

The genomic classifier divided the TCGA cohort to two subtypes (A-C): high-risk group (n=69) and low risk group (n=13). Chi-squared tests indicated that high-risk group has a higher tumor stage, lymph node stage, more frequent new neoplasm event, lower hemoglobin level and relative higher genomic alteration frequency ( $p < 0.05$ , D-H). GSEA indicated that compared with low-risk group, gene expression pattern of high-risk group were significantly enriched in E2F targets, G2M checkpoint, Myc targets and other pathways (I). To our interest, the gene expression pattern of high-risk group also enriched in mTORC1 signaling pathway (J), which may shed light on the accurate use of mTOR Inhibitors such as everolimus.



## Figure 6

As described in the methods, TPX2 expression level was used to replace the genomic classifier and it could fully match the clusters (A-B). 92 patients with histopathologic confirmation of sPRCC2 (positive staining of FH) was enrolled (C-D) and pictures of TPX2 expression (IHC, low and high) were shown (E-F). Samples with Low TPX2 expression and high TPX2 expression were defined as low (N=49) and high-risk group (N=43). Survival analysis indicated that both the overall survival ( $p < 0.0001$ , HR=3.361, G) and progression free survival ( $p < 0.0001$ , HR=4.209, H) of high-risk group was significantly worse than the low risk group. Although statistically not significant, patients in high-risk group (N=7) has a better progression free survival than the low-risk group (N=6) after receiving everolimus as first-line therapy (I). Retrospective analysis also indicated that everolimus showed better effect (J) in high-risk group than sunitinib (N=5). In summary, everolimus showed stronger drug effect in high-risk group sPRCC2 than that in low-risk group (overall response rate: 28.6%, 16.7%) and everolimus is better than sunitinib (overall response rate: 28.6%, 20%) for high-risk group (better overall response and reduction of the target lesion, K).

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.tif](#)
- [TableS1.xlsx](#)
- [TableS2.xlsx](#)
- [TableS3.xlsx](#)
- [TableS4.xlsx](#)