



Development and Clinical Validation of a Seven-Gene Prognostic Signature Based on Multiple Machine Learning Algorithms in Kidney Cancer

Cell Transplantation
Volume 30: 1–17
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0963689720969176
journals.sagepub.com/home/ctj


Mi Tian¹, Tao Wang², and Peng Wang³ 

Abstract

About a third of patients with kidney cancer experience recurrence or cancer-related progression. Clinically, kidney cancer prognoses may be quite different, even in patients with kidney cancer at the same clinical stage. Therefore, there is an urgent need to screen for kidney cancer prognosis biomarkers. Differentially expressed genes (DEGs) were identified using kidney cancer RNA sequencing data from the Gene Expression Omnibus (GEO) database. Biomarkers were screened using random forest (RF) and support vector machine (SVM) models, and a multigene signature was constructed using the least absolute shrinkage and selection operator (LASSO) regression analysis. Univariate and multivariate Cox regression analyses were performed to explore the relationships between clinical features and prognosis. Finally, the reliability and clinical applicability of the model were validated, and relationships with biological pathways were identified. Western blots were also performed to evaluate gene expression. A total of 50 DEGs were obtained by intersecting the RF and SVM models. A seven-gene signature (RNASET2, EZH2, FXYD5, KIF18A, NAT8, CDCA7, and WNT7B) was constructed by LASSO regression. Univariate and multivariate Cox regression analyses showed that the seven-gene signature was an independent prognostic factor for kidney cancer. Finally, a predictive nomogram was established in The Cancer Genome Atlas (TCGA) cohort and validated internally. In tumor tissue, RNASET2 and FXYD5 were highly expressed and NAT8 was lowly expressed at the protein and transcription levels. This model could complement the clinicopathological characteristics of kidney cancer and promote the personalized management of patients with kidney cancer.

Keywords

kidney cancer, SVM, RF, LASSO, gene signature, prognosis

Introduction

Kidney cancer is a malignant tumor that originates in renal, parenchymal, and urinary epithelial cells, and there are approximately 208,500 new cases each year worldwide¹. Despite recent advances in clinical chemotherapy and surgical techniques, the prognosis of advanced kidney cancer is still poor, with a median survival of approximately 13 months². About 30% of patients with kidney cancer experience recurrence or cancer-related progression^{3,4}. Nearly 20% of kidney cancer cases have progressed to advanced stages by the time of diagnosis⁵. At present, the most valuable factor for kidney cancer prognosis is the pathological stage after surgery. The earlier the pathological stage, the higher the 10-year survival rate^{6–8}. If lymph node metastasis has occurred by the time of diagnosis, the prognosis is often not positive⁹. Tumor size, histological type, and the cell

cycle are also independent factors affecting prognosis^{10–12}. Although these indicators play important roles in determining prognosis, patients with kidney cancer at the same

¹ Department of Nephrology, Shengjing Hospital of China Medical University, Shenyang, China

² Department of Pathology, Shenyang KingMed Center for Clinical Laboratory Co, Ltd, Shenyang, China

³ Department of Ultrasound, Shengjing Hospital of China Medical University, Shenyang, China

Submitted: June 21, 2020. Revised: September 19, 2020. Accepted: October 7, 2020.

Corresponding Author:

Peng Wang, Department of Ultrasound, Shengjing Hospital of China Medical University, Shenyang, Liaoning 110004, PR China.
Email: wangpcmu@126.com



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

clinical stage could have quite different prognoses. With the rapid development of tumor molecular biology in recent years, genomic studies have shown the increasing importance of RNA. Therefore, there is an urgent need to identify the RNAs that may improve the clinical outcomes of patients with kidney cancer¹³. However, there are few specific biomarkers that show therapeutic effects, and prognostic factors are also important for treatment. Therefore, the molecular screening of kidney cancer biomarkers is necessary to improve kidney cancer prognosis and reduce mortality.

Messenger RNAs (mRNAs) are single-stranded ribonucleic acids transcribed from single strands of DNA and they carry genetic information to guide protein synthesis¹⁴. Most mRNAs exist in the cytoplasm of prokaryotes and eukaryotes and in certain organelles of eukaryotic cells. They serve as templates for protein synthesis on ribosomes and determine the amino acid sequences of peptide chains^{15,16}. The expression level and methylation of mRNA are closely related to the processes of tumorigenesis, proliferation, and invasion^{17–19}. In addition, mRNAs are closely associated with the prognoses of various tumors and can be used as reliable biomarkers for prognosis^{20–22}. Several genes with potential clinically and statistically significant prognostic value have been identified in the whole blood expression profiles of patients with clear cell renal cell carcinoma²³. Other studies have also identified individual biomarkers related to kidney cancer prognosis^{24,25}. Integrating multiple mRNAs into a single model has more reliable prediction and prognostic value than screening only a single mRNA as a biomarker²⁶. Karakiewicz et al. proposed a preoperative prognostic model of kidney cancer. For patients with kidney cancer undergoing nephrectomy, the model's predictive ability showed high accuracy²⁷. Heng et al. proposed a well-known prognostic model for patients with metastatic kidney cancer that has played an important role in selecting suitable participants in many clinical trials²⁸. Further, the five-mRNA model constructed by Gao et al. using the RNA-seq dataset from The Cancer Genome Atlas (TCGA) is considered a potential prognostic biomarker for papillary kidney cancer²⁹.

The purpose of this study was to use the RNA sequencing data from TCGA and the Gene Expression Omnibus (GEO) pertaining to kidney cancer to explore the expression differences between renal cancer and adjacent tissue and to identify potential prognostic biomarkers. Based on survival analysis, a seven-gene prognostic model, including RNA-SET2, EZH2, FXYD5, KIF18A, NAT8, CDCA7, and WNT7B, was established, and the relationships with prognosis and clinical features were analyzed. Finally, a predictive nomogram was established in the TCGA cohort and validated internally. At the same time, we used experiments and external cohorts to verify the mRNA and protein expression levels of these seven genes. The prognostic model and nomogram may help guide the evaluation of prognostic status for patients with kidney cancer.

Materials and Methods

Data Downloading and Preprocessing

In the GEO database, the chip data of clear cell renal cell carcinoma tissue samples were collected using the search term “kidney cancer” as the keyword, with the search scope limited to “Homo sapiens” (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53757>). The dataset included a cancer tissue sample ($n = 77$) and a cancer-adjacent sample as the control ($n = 77$). After the chip data were cleaned and compared, 22,880 genes were obtained. All genes and samples were found to contain no missing values, so they could proceed to the next step. As shown in **Fig. S1**, the corrected expression values of the samples were all on the same level, indicating that they were well homogenized.

Sample Grouping

Fragments per kilobase million (FPKM) data and clinical information pertaining to RNA expression in renal cell carcinoma were downloaded from the TCGA database. Data with no clinical prognosis information or a gene expression level <1 were excluded. The *caret* package in R was used to randomly divide the cohort with a ratio of 7.5:2.5, with 75% of the data used for training and 25% of the data used for validation.

Construction of Gene Signature by Integrating Multiple Machine Learning Algorithms

In the random forest (RF) model, a fivefold cross-validation method was adopted to divide the training set and the validation set, iterate on the number of variables tried at each split and ntree (500 to 1000), respectively, and finally find that when the number of variables tried at each split = 58 and ntree = 500, the out-of-bag error (OOB) is the lowest (2.78%).

In the support vector machine (SVM), the optimal variables are identified by deleting the feature vectors generated by SVM, which includes the following steps. (1) Fitting a linear SVM model. (2) Sorted according to the weight of features in each SVM model. (3) Eliminate variables with low weights. The fivefold cross-validation method was used to split the data and assign random numbers. When halve above was set to 100, the order of feature vector and mean value was obtained, and AverageRank > 2500 was taken as threshold.

The least absolute shrinkage and selection operator (LASSO) method is a compression estimation. It obtains a more refined model by constructing a penalty function, which makes it compress some coefficients and set some coefficients to zero. Therefore, the advantage of subset shrinkage is retained. It is a biased estimation for processing data with multicollinearity, which can realize variable selection while estimating parameters, and better solve the multicollinearity problem in regression analysis. We use the

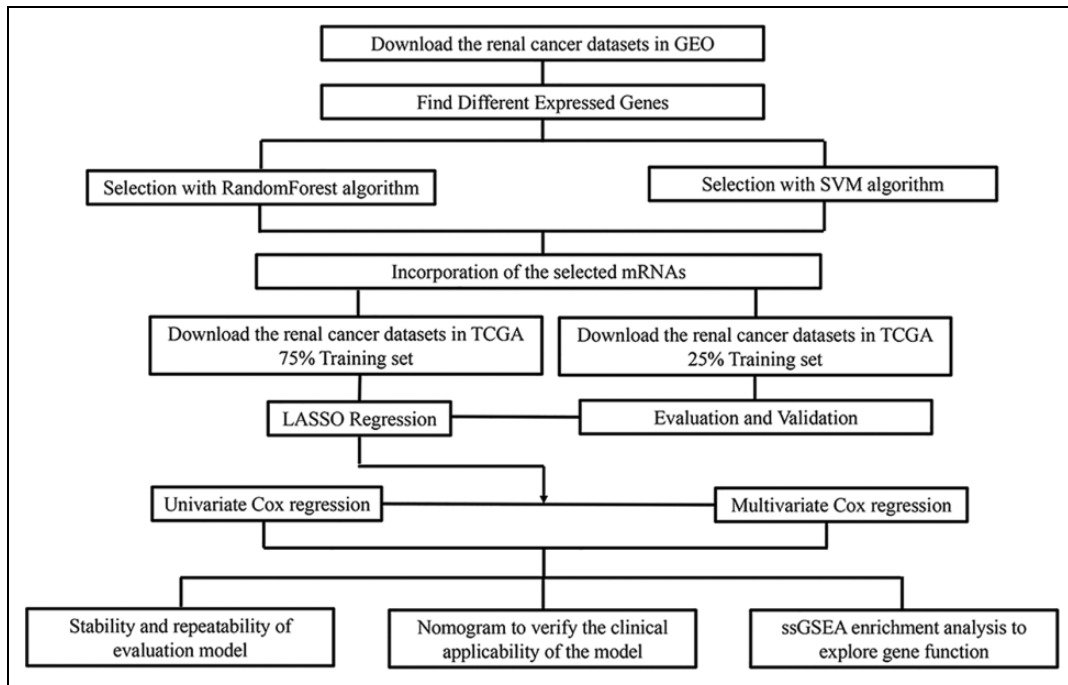


Fig. 1. Flowchart of this study.

glmnet package in R to perform LASSO analysis on the RNAs selected by the RF and SVM models.

External Validation of Proteins and Transcription Levels of the Seven-Gene Signature

The Human Protein Atlas (HPA) provides information on the tissue and cell distribution of 26,000 human proteins. It mainly uses specific antibodies to study the protein expression in cell lines, normal tissue, and tumor tissue. The expression of seven genes (RNASET2, EZH2, FXD5, KIF18A, NAT8, CDCA7, and WNT7B) was explored in normal and tumor tissue. The expression of these seven genes was explored in kidney cancer and normal tissue in the GSE105288 and GSE106771 series of the GEO. Box-plots for gene expression were drawn.

Genetic Alterations of the Seven Predictive Genes

The cBioPortal database integrates genomic data, including somatic mutations, DNA copy-number alterations, mRNA and microRNA (miRNA) expression, DNA methylation, protein enrichment, and phosphorylated protein enrichment. Clear Cell Renal Cell Carcinoma (DFCI, Science 2019), Kidney Chromophobe (TCGA, Cancer Cell 2014), Kidney Chromophobe (TCGA, Firehose Legacy), Kidney Chromophobe (TCGA, Pan Cancer Atlas), and Kidney Renal Clear Cell Carcinoma (BGI, Nat Genet 2012) datasets were collected.

Western Blotting

Western blotting was performed according to standard protocols. We used primary antibodies raised against glyceraldehyde 3-phosphate dehydrogenase (Santa Cruz Biotechnology, CA, USA); WNT7B, CDCA7, KIF18A, and EZH2 (Cell Signaling Technology, MA, USA); and FXD5, NAT10, and RNASET2 (Proteintech, China). Goat anti-mouse and antirabbit antibodies conjugated with horseradish peroxidase were used as secondary antibodies (Jackson ImmunoResearch, West Grove, PA, USA), and we detected the blots using enhanced chemiluminescence (Dura, Pierce, NJ, USA).

RNA Extraction and Real-Time Polymerase Chain Reaction (PCR) Assay

Total RNA was extracted using TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's protocol, and it was reverse-transcribed into complementary DNA (cDNA) using a Superscript Reverse Transcriptase Kit (Transgene, France). A Super SYBR Green Kit (Transgene, France) was used to perform real-time PCR using the ABI 7300 real-time PCR system (Applied Biosystems, Foster City, CA, USA). The primer pairs were: (1) WNT7B forward: CACAGAACTTTCGCAAGTGG, WNT7B reverse: GTACTGGCACTCGTTGATGC; (2) CDCA7 forward: TTGGTCTTCGAGTAGCCTTCA, CDCA7 reverse: GTGCGCTAGAAAACAAGTCT; (3) KIF18A forward: TGCTGGGAAGACCCACACTAT, KIF18A reverse: GCTGGTGTAAGTAAGTCCATGA; (4) EZH2

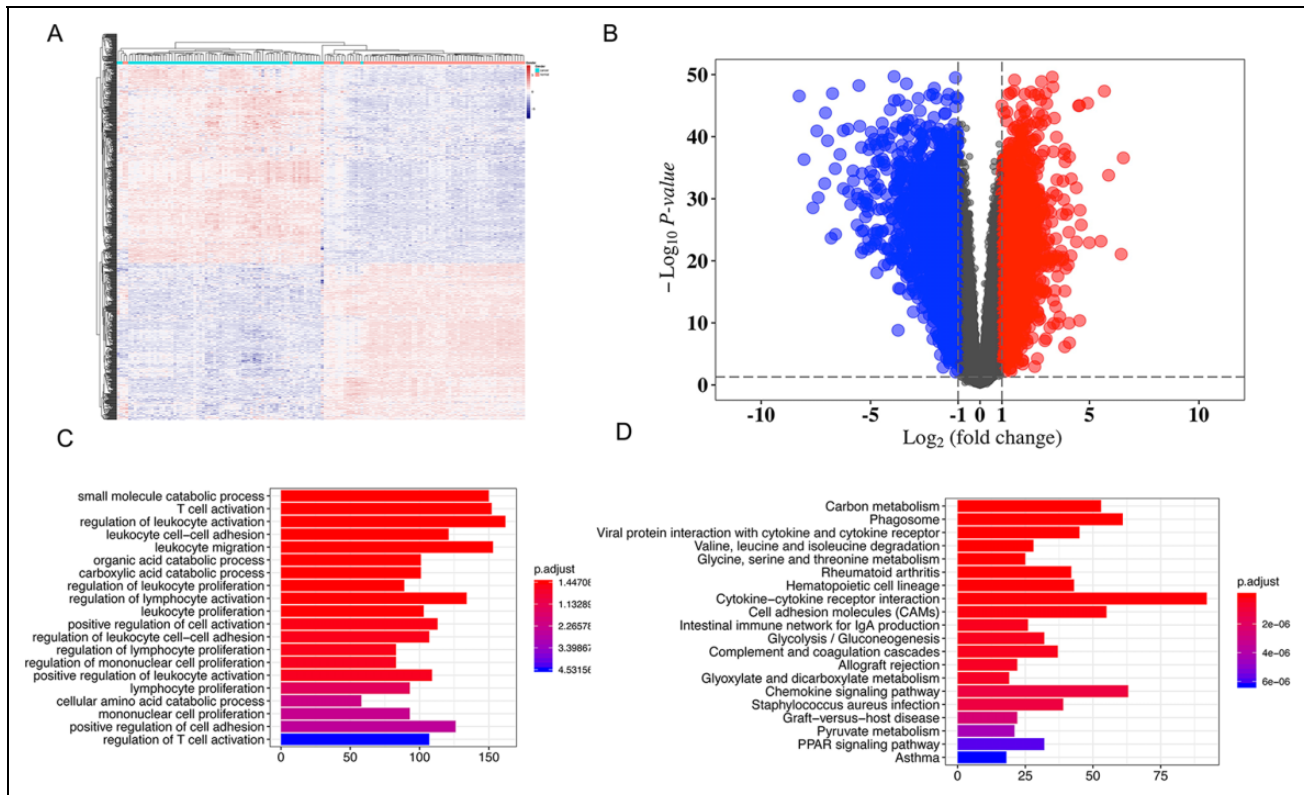


Fig. 2. Identification of Differentially Expressed Genes and functional analysis. (A) Heat map of differential genes. Genes highly expressed in the samples are in red, while the others are in blue. The expression of genes in the cancer group and the paracancerous group were quantified into two obvious panels. (B) Volcano map of differential genes, the abscissa is $\log_2 FC$, the ordinate is $-\log_{10} P$ value, and different points represent different RNAs. Genes highly expressed in the samples are in red, while the others are in blue. (C) Gene Ontology enrichment analysis. (D) Kyoto Encyclopedia of Genes and Genomes enrichment analysis.

forward: AATCAGAGTACATGCGACTGAGA, EZH2 reverse: GCTGTATCCTTCGCTGTTTCC; (5) FXD5 forward: AGTGGTCATCCTCCTACGGAC, FXD5 reverse: TGTACCTGGAATGCACATCCAT; (6) NAT10 forward: ATAGCAGCCACAAACATTTCGC, NAT10 reverse: ACA-CACATGCCGAAGGTATTG; and (7) RNASET2 forward: GCGAGAAAATTCAAACGACTGT, RNASET2 reverse: CCTTCACTTTTATCGGGCCATAG.

Results

Data Analysis Flowchart

To make the research easier for readers to understand, we drew a flowchart of the methodology (Fig. 1).

Differentially Expressed Genes (DEG) Identification

The GSE53757 dataset was used for different analysis and screening for marker genes. A total of 3471 RNAs were screened using the *limma* package (adj. P value < 0.05 and $|\log_2 FC| > 1$). The heat map is shown in Fig. 2A. The blue bar represents the cancer group, and the red bar represents the normal group. Gene expression in cancer and

paracancerous groups was quantified into two obvious panels, indicating that the differential genes screened had obvious differences in expression between the groups. The $\log_2 FC$ and $-\log_{10} P$ value of the differential genes were plotted into a volcano map, as shown in Fig. 2B.

Functional Analysis of DEGs

Further analysis with the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) functional enrichment of these DEGs through the R package *clusterProfiler* was performed to explore their biological functions, and the threshold was set at $P < 0.05$. As shown in Fig. 2C, D, the DEGs were enriched in many important cancer-related pathways, such as the T-cell activation pathway.

Construction of Seven-Gene Signature

According to the Materials & Methods section, 556 RNAs were identified as candidate genes in the RF model. Four hundred and twenty-two RNAs were selected as candidate genes in the SVM model. The marker genes obtained from the RF and SVM models were intersected, and 50 marker genes were identified for subsequent analysis (Fig. 3A, B).

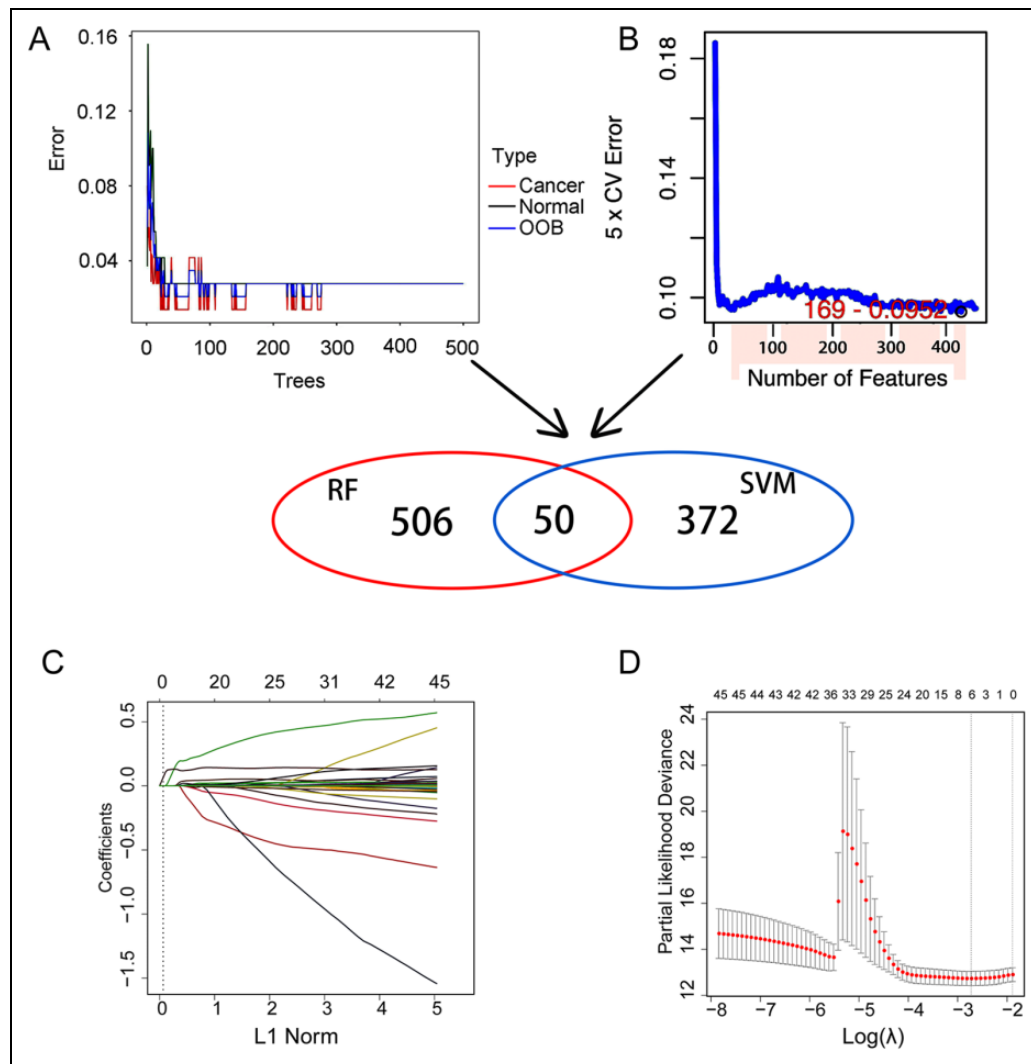


Fig. 3. Construction of multigene signatures. (A) Error graph of the random forest (RF) models. (B) Support vector machine (SVM) models. (C) The trajectory of each independent variable, the horizontal axis represents the log value of the independent variable lambda, and the vertical axis represents the coefficient of the independent variable. (D) The CI at different lambda values.

These genes are closely related to the occurrence and development of kidney cancer, but too many genes are not conducive to clinical testing and will also increase the burden of patients. Therefore, we further reduce the number of genes.

The *glmnet* package was used to perform LASSO cox analysis on these candidate genes. First, the change trajectory of each independent variable was analyzed as shown in **Fig. 3C**. It can be seen that as the lambda gradually decreases, the number of independent variable coefficient tends to 0 is gradually increasing. We use 10-fold cross-validation to analyze the CI under each lambda as shown in **Fig. 3D**. It can be seen that the model reaches optimal when $\ln(\lambda) = -2.56$. We select the seven genes (RNASET2, EZH2, FXYD5, KIF18A, NAT8, CDCA7, and WNT7B) at this time as the target genes to construct the seven-gene signature. The risk scores were calculated with the

expression profile of the seven genes: $(0.00134 * \text{expression level of RNASET2}) + (0.00968 * \text{expression level of EZH2}) + (0.00261 * \text{expression level of FXYD5}) + (0.17314 * \text{expression level of KIF18A}) + (0.00001 * \text{expression level of NAT8}) + (0.12862 * \text{expression level of CDCA7}) + (0.00001 * \text{expression level of WNT7B})$.

The seven-gene signature is different from the existing stage system, which evaluates the prognosis of patients from the perspective of molecular biology. In clinical use, only the expression levels of these seven genes need to be detected, according to the risk score formula.

The distribution of risk score Z correction values was plotted. The results showed that patients in the high-expression group (red) had significantly higher risk scores than patients in the low-expression group (blue), and the density distribution chart was shifted to the right (**Fig. 4A**). The seven-gene signature risk scores of each patient

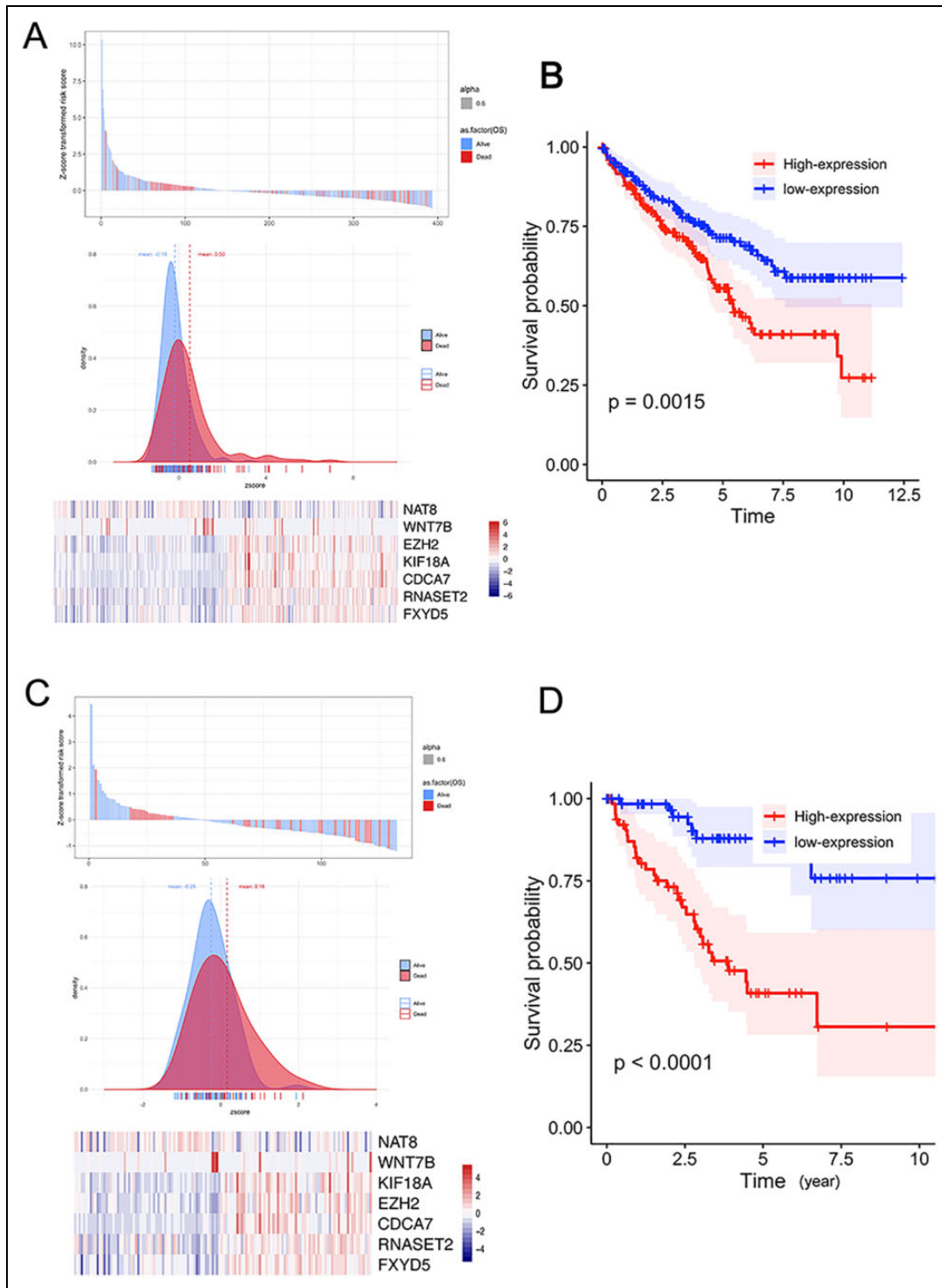


Fig. 4. The performance of the seven-mRNA model. (A) The Z correction value of risk score distribution, the risk score density distribution, and the seven gene expression heat maps. (B) Survival analysis. (C) The Z correction value of risk score distribution, risk score density distribution, and seven gene expression heat maps in the validation set. (D) Survival analysis.

were calculated in the TCGA training set, and the median risk score was used as the cutoff. The patients were divided into the low-risk group ($n = 262$) and high-risk group ($n = 263$). Patients in the high-risk group had significantly shorter

overall survival (OS) than those in the low-risk group (**Fig. 4B**). The red line represents the high-risk group, and the blue line represents the low-risk group. The P value between groups was 0.0015.

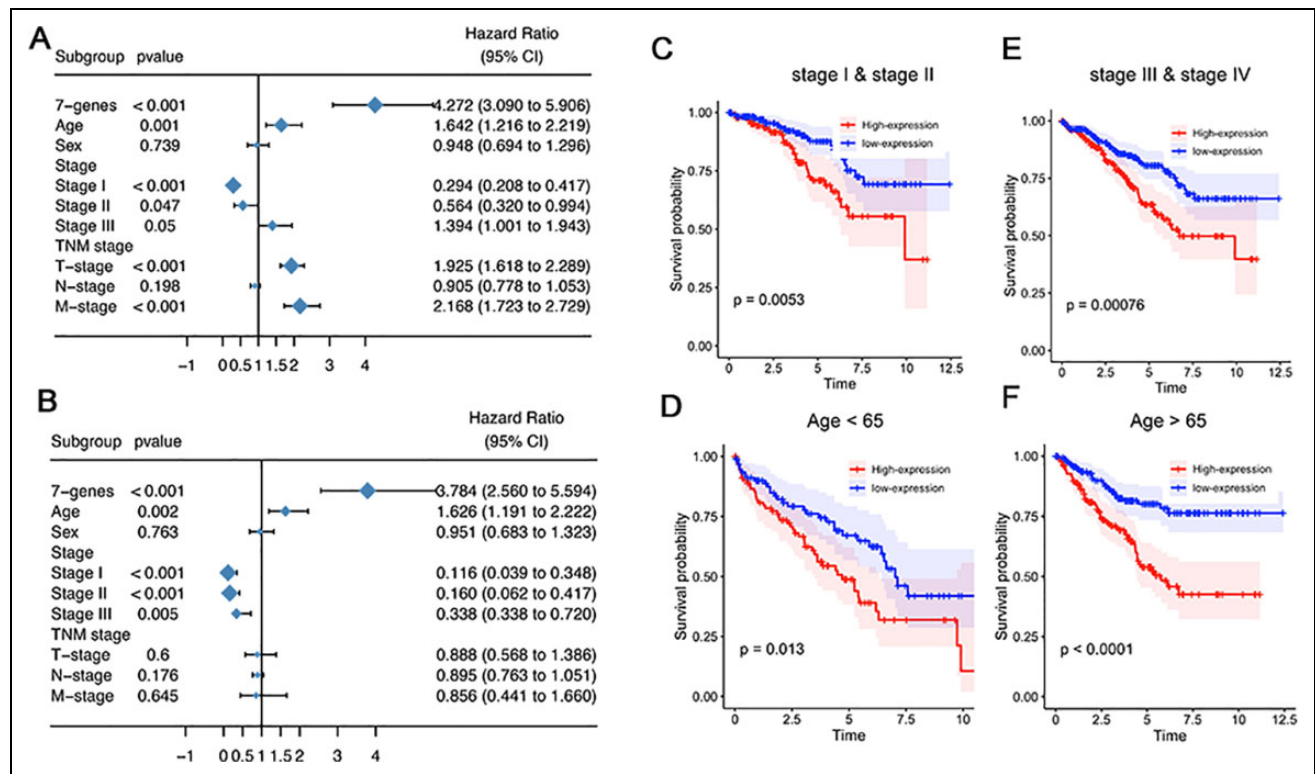


Fig. 5. Analysis of univariable and multivariable Cox regression of seven-gene signature. (A) Forest plots of univariable Cox model; (B) forest plot of the multivariable Cox model. (C-F) Survival analysis of different subgroups assessing the independence of seven-gene signatures.

The TCGA validation set was used for verification. The P value between groups was < 0.05 , and the distribution map and density distribution of risk score Z correction values were basically consistent with those of the training set. **Fig. 4C-D** shows that this model had a quite good repeatability.

Using this seven-gene based risk score prognostic model, the patient's prognostic score was evaluated. When the patient's risk score is greater than 0, the patient is at high-risk. The clinician can change the patient's treatment plan according to the predicted results of the model to realize the individualized treatment of patients. Strategies should be developed to prevent or detect recurrence early in high-risk groups. Therefore, high-risk groups should be followed more frequently.

Univariable and Multivariable Cox Regression Analysis of the Seven-Gene Signature

To prove the prognostic value of seven-gene signature in kidney cancer patients, the univariable and multivariable Cox regression analyses were performed on seven-gene risk score, age; sex, clinical stage, and pathological tumor, node, metastasis (TNM) stage.

Hazard ratios and 95% CIs were calculated. Results with a P value < 0.05 were considered statistically significant.

The results are shown in **Fig. 5A, B**. The seven-gene risk score was significantly associated with prognosis in both the univariable and multivariable analyses, and it was an independent risk factor for kidney cancer prognosis ($P < 0.001$, hazard ratio = 3.784).

Both age and stage had $P < 0.05$, indicating that they were important prognostic factors, and there was no significant difference in the TNM stage or sex.

To eliminate the confounding factors of age and clinical stage, the sample was divided into early stage (stage I, stage II), advanced stage (stage III, stage IV), younger age (< 65 years), and older age (≥ 65 years) groups. The results showed that after excluding age and stage, there were still significant differences between the two groups ($P < 0.05$), and the seven-gene signature distinguished the high-expression group from the low-expression group (**Fig. 5C-F**).

Evaluation of the Clinical Applicability of the Seven-Gene Signature

In order to determine the patient's disease progression and be able to make a personalized diagnosis of the patient, a risk prediction nomogram integrating the seven-gene signature, age, stage, TNM stage, and sex was plotted (**Fig. 6A**). **Fig. 6B** shows that the calibration of the nomogram worked well compared with the actual model. Decline curve analysis was

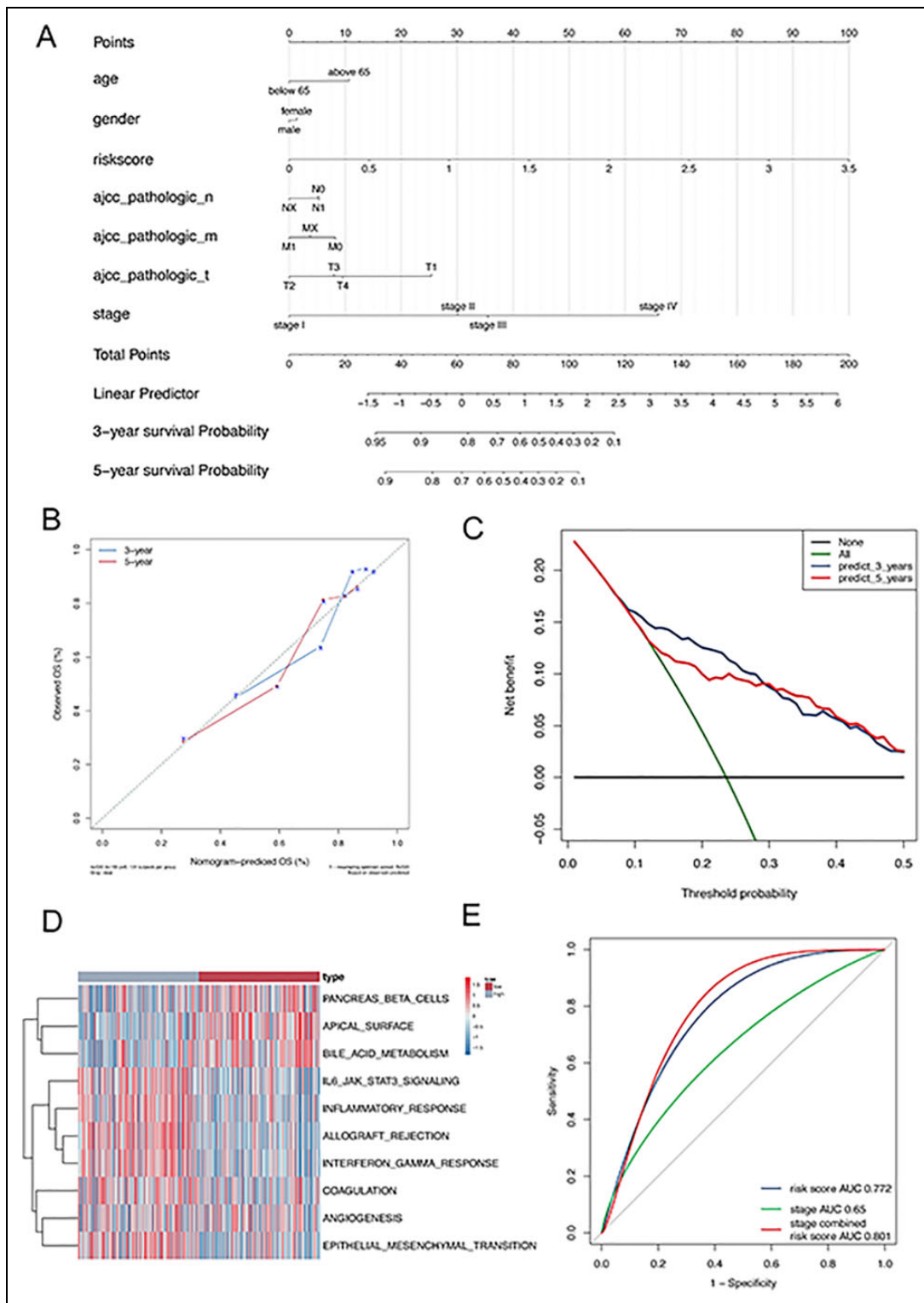


Fig. 6. Evaluation of the clinical applicability of the seven-gene signature. (A) The nomogram for predicting the proportion of patients with 3-year OS and 5-year OS. (B) The calibration plots for predicting patient 3-year OS and 5-year OS. Nomogram-predicted probability of survival is plotted on the x-axis; actual survival is plotted on the y-axis. (C) DCA for assessment of the clinical utility of the nomogram. The x-axis represents the percentage of threshold probability, and the y-axis represents the net benefit. (D) Pathway profiles. Rows represent pathways, and columns represent patients. Each grid represents a score of pathway activity calculated by single-sample GSEA. The upper horizontal bar marked the information related to every patient, including its risk group (ranked from low to high). (E) ROC analysis of the sensitivity and specificity of the survival prediction by the seven-gene risk score. DCA: decision curve analysis; GSEA: gene set enrichment analysis; OS: overall survival; ROC: receiver operating characteristic.

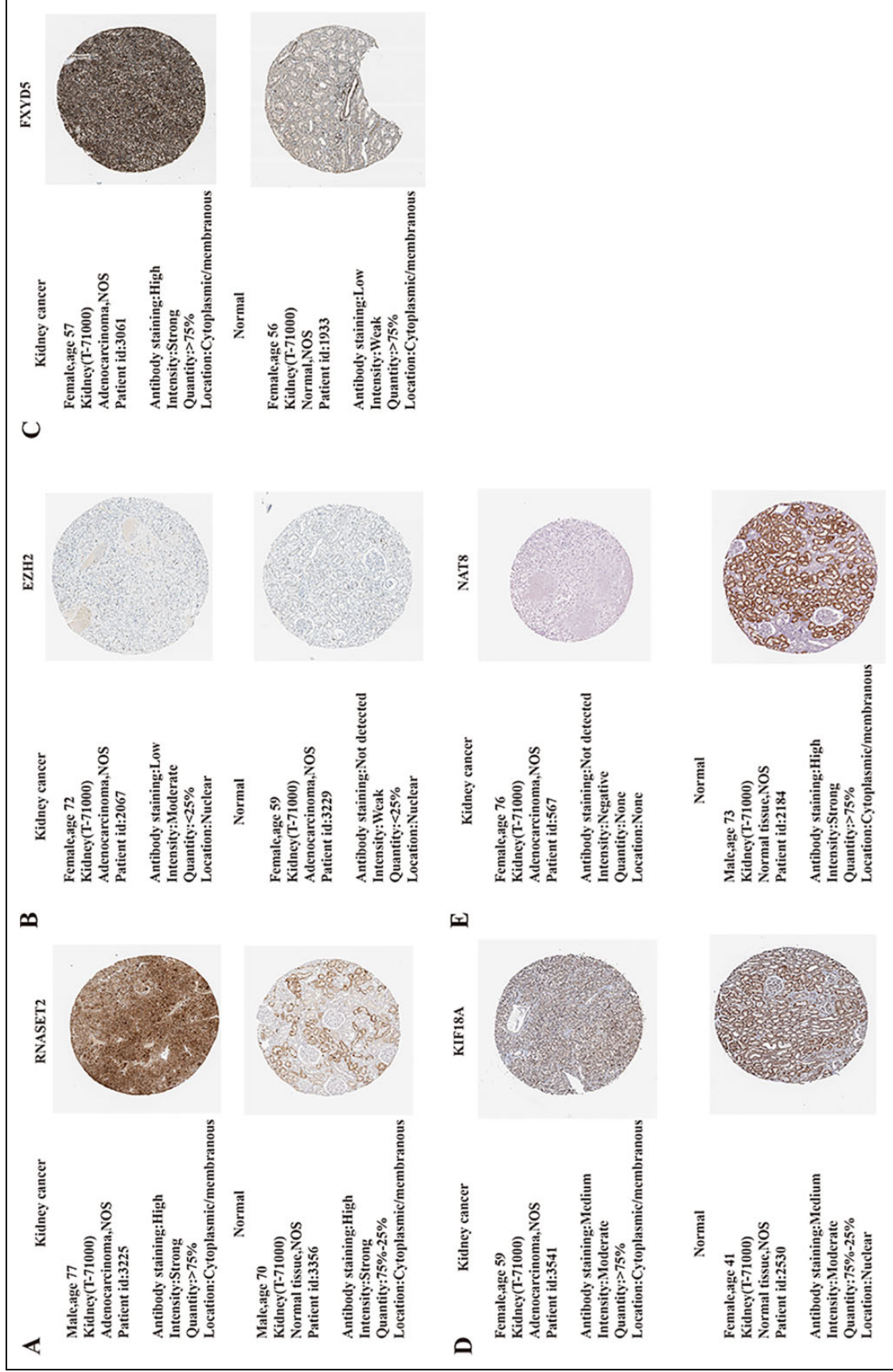


Fig. 7. Expression of seven-gene signature in the Human Protein Atlas protein database.

used to evaluate the clinical applicability of the nomogram. The results showed that our line chart model had a better net benefit (**Fig. 6C**). The receiver operating characteristic curve is shown in **Fig. 6E**. The model with the stage alone had a minimum area under the curve (AUC) of 0.65. The risk score had an AUC of 0.772. A combination of risk score and stage had the highest AUC (0.801).

The low-expression and high-expression groups were used for classification, and enrichment analysis was performed using a single-sample gene set enrichment analysis (ssGSEA), as shown in **Fig. 6D**. Some biochemical-related pathways had higher enrichment scores in the low-expression group, while tumorigenic-related pathways had higher enrichment scores in the high-expression group, which justified the previous results.

Validation of the Expression of the Seven-Genes Signature

The HPA database was used to analyze the protein expression levels of the seven genes. Among them, EZH2 and KIF18A were not significantly different between cancer and normal tissue, RNASET2 and FXYD5 were relatively highly expressed in cancer, NAT8 was relatively lowly expressed in cancer, and CDCA7 and WNT7B had no corresponding protein expression (**Fig. 7A-E**).

We then measured the expression levels of the seven genes in three pairs of kidney cancer and normal tissue. We found that compared with normal controls, RNASET2 and FXYD5 were significantly highly expressed in cancer, and NAT8 was relatively lowly expressed in cancer. There was no difference in expression between tumor tissue and adjacent tissue in EZH2, KLF18A, CDCA7, and WNT7B, and the experimental results were almost consistent with our data analysis (**Fig. 8**).

Genetic Alterations of the 7 Genes

The mutations of the seven genes were analyzed in the cBioPortal database. The gene with the highest mutation proportion was EZH2, accounting for 0.8%, and its mutation types were amplification and point mutation (**Fig. 9A-H**).

Discussion

The expression level of mRNA might be related to the development of various types of tumors³⁰. Some mRNAs are considered potential biomarkers for predicting kidney cancer prognosis, such as the fat mass and obesity-associated (FTO), ferritin heavy chain (FTH1), and thioredoxin domain-containing 5 (TXNDC5) genes, and so on³¹. However, existing prognostic biomarkers for clinical application still have great limitations, such as insufficient samples, too few mRNAs, or lack of independent validation³². The reliability and usefulness of biomarkers require further verification. To build a more reliable mRNA prognostic model,

existing gene expression databases were searched to screen for mRNAs with prognostic significance. In this research, two models (RF and SVM) were used. RF is a tree-based classification algorithm that can be applied for classification and regression³³. SVM is a powerful classification tool based on statistical learning theory. It can be used to create a boundary between two categories, and so it can predict targets based on one or more feature vectors; it is thus suitable for performing classification and regression and probability estimations³⁴. These two algorithms have played an increasingly important role in the detection of cancer, the exploration of changes in specific functions of different cancer types, and so on³⁵. In this study, by exploring the correlation between mRNA expression profiles in the GSE53757 dataset and clinical kidney cancer prognosis, a seven-gene independent prognostic model significantly correlating with kidney cancer prognosis was constructed.

First, kidney cancer differential genes were screened from the dataset. To further understand the biological functions of these differential genes, functional enrichment analysis was performed. Biological functions are mainly enriched in terms of small molecule catabolic processes, T-cell activation, and other aspects. Small molecule inhibitors and hormone-like drugs have been used in the treatment of cancer³⁶. Small molecule metabolites can be used as potential biomarkers of liver cancer and better distinguish liver cancer from other bile duct diseases with bile duct tumor thrombosis³⁷. The activation of effector CD8 T cells can initiate an immune response that is positively associated with breast cancer survival and be used as a marker for personalized immunotherapy³⁸. Biological functions are also enriched in various biological behaviors of leukocytes, such as regulation of leukocyte activation, leukocyte cell-cell adhesion, leukocyte migration, regulation of leukocyte proliferation, and leukocyte proliferation. Neutrophils account for 50% to 70% of all white blood cells and can reflect the state of inflammation, which is an important sign of cancer, in the host^{39,40}. Neutrophils are involved in different stages of the cancer process, including tumorigenesis, growth, proliferation, and metastasis, and they can be used as biomarkers or therapeutic targets for prognosis^{41,42}. Neutrophils can promote tumor proliferation by weakening the body's immune system⁴³. They can also promote tumorigenesis by releasing reactive oxygen species or proteases⁴⁴ and stimulate metastasis and tumor spread by inhibiting natural killing functions and promoting tumor cell exudation^{45,46}. Gene-related pathways are mainly concentrated in the carbon metabolism, phagosome, cytokine-cytokine receptor interaction, cell adhesion molecule, and chemokine signaling pathways. Cells need a carbon atom unit for nucleotide synthesis. These pathways are closely related to the high proliferation rate of cancer cells. Drugs targeting carbon metabolism have been employed in clinical cancer treatment⁴⁷. Delayed phagosome maturation retains antigenic peptides for presentation to T cells and initiates adaptive immune responses, and adaptive immune resistance is an important process for cancer to

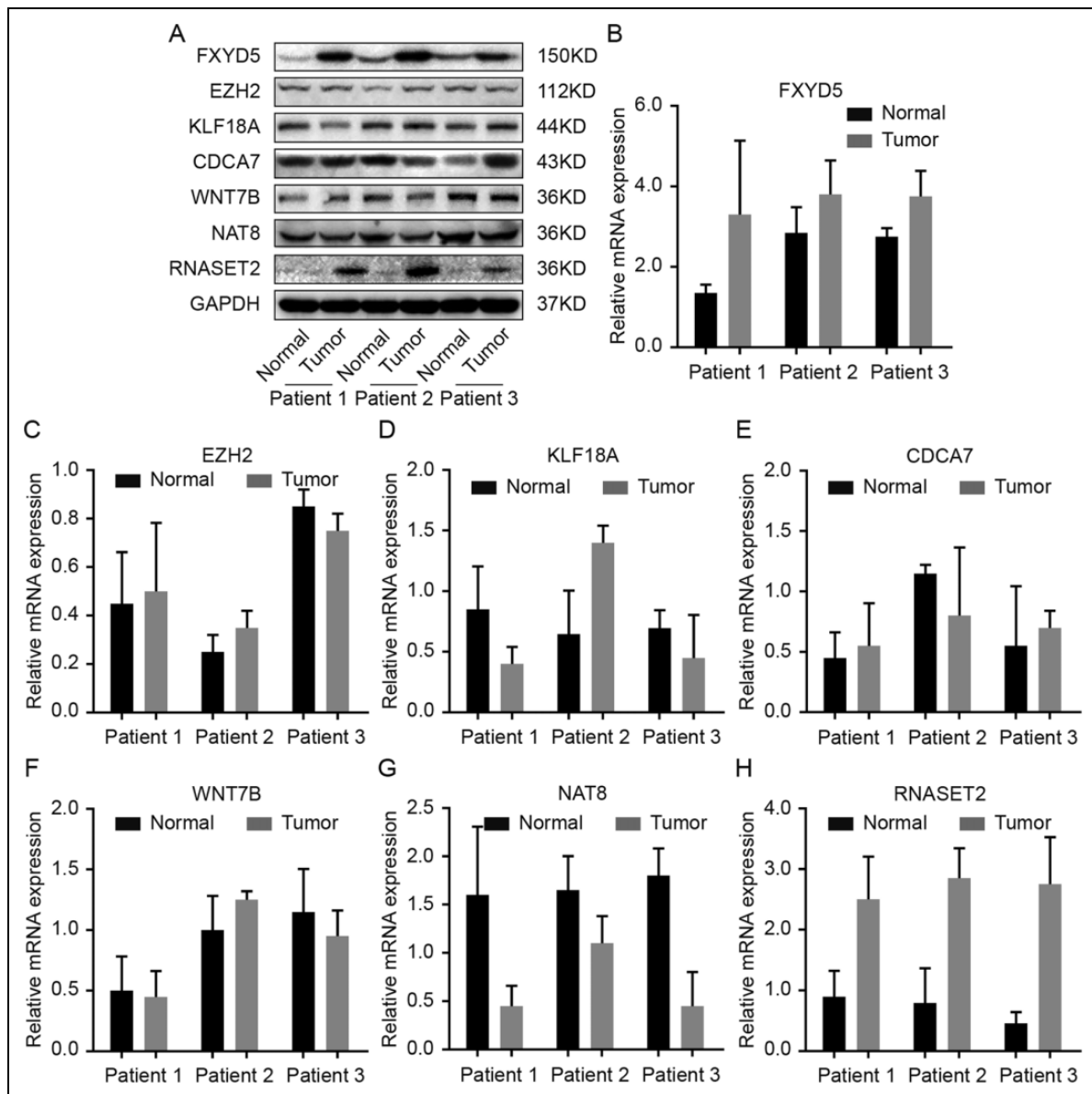


Fig. 8. Protein and mRNA expression of seven-gene signature in three pairs of renal cell carcinoma and paracancerous tissues.

evade immunity by changing its phenotype^{48,49}. Cytokines play an immunoregulatory role and are essential for the development of human biology and disease⁵⁰. For example, in the interactions of many cytokines with cytokine receptors, interleukin (IL)-21 exerts effective antitumor effects⁵¹. Chemokines are a family of small molecule cytokines whose main function is to attract immune cells to the corresponding site of inflammation. The chemokine signaling pathway is closely related to the growth, proliferation, invasion, and metastasis of cancer cells⁵²⁻⁵⁵. Cell adhesion molecules are also involved in cancer progression and metastasis, and they can be potential predictors of tumor recurrence⁵⁶⁻⁵⁸. The pathways enriched by GO and KEGG are involved in the genesis and development of tumors, which further validates the correlation between the differential genes screened and

tumor biological behavior, suggesting that these differential genes might be related to prognosis.

To further identify the genes most closely related to kidney cancer prognosis, RF and SVM models were constructed, and seven hub mRNAs were screened out by lasso cox analysis. Finally, we determined a seven-gene prognostic signature, which is different from the existing stage system, which evaluates the prognosis of patients from the perspective of molecular biology.

Previous studies have shown that these genes are involved in tumor development. The RNASET2 gene is the only member of the T2 extracellular ribonuclease family that has been localized in humans. It was found to be rearranged in a variety of cancers and is often considered a tumor suppressor gene⁵⁹⁻⁶², yet there are currently few studies of this gene in

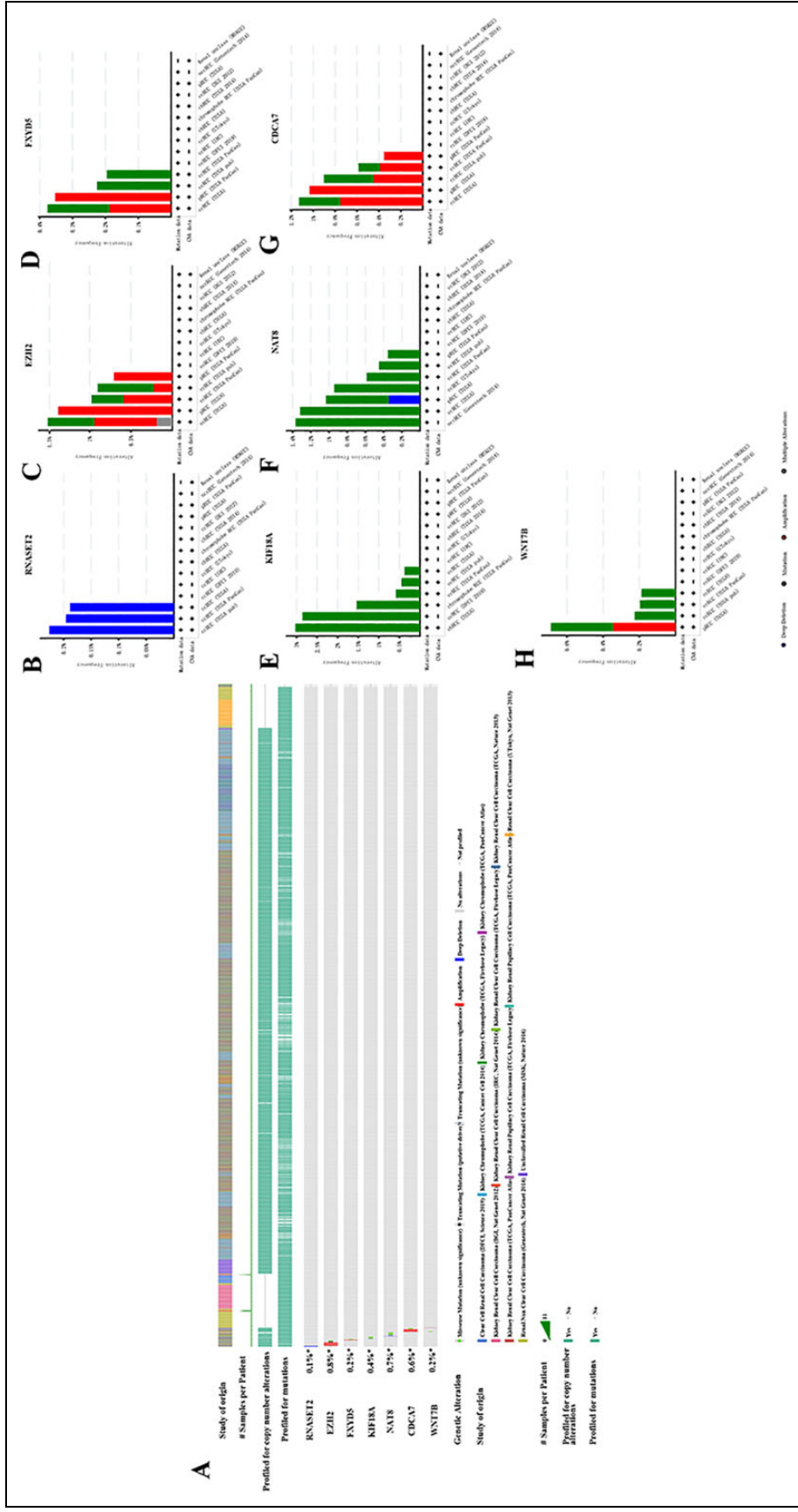


Fig. 9. Genetic alterations of the seven genes. (A) Oncoprint display of mutation distribution of seven genes. (B) Histogram display of mutation distribution of seven genes.

relation to kidney cancer. EZH2 is a pleiotropic molecule, and its basic function is to participate in epigenetic gene suppression as a component of the poly-comb repressive complex 2 (PRC2)⁶³. Phosphorylated EZH2 is also a co-activator of key transcription factors⁶⁴. EZH2 has been widely studied in cancer prognosis, and its expression level is closely related to the prognosis of tumors such as lung adenocarcinoma, breast cancer, glioma, and intrahepatic cholangiocarcinoma^{65–68}. The expression level of EZH2 is elevated in BAP1 mutant renal carcinoma, and it is related to the poor prognosis of renal carcinoma⁶⁹. FXYD5 is a type I membrane protein and an important member of the FXYD family. FXYD5 is an aggressive biomarker of endometrial cancer and closely related to the prognosis of ovarian cancer, but its biological role in kidney cancer is not clear^{70–72}. KIF18A is a member of the kinesin superfamily, and its overexpression is associated with the poor prognosis of primary hepatocellular carcinoma^{73,74}. KIF18A can also predict the OS rate of patients with papillary renal cell carcinoma and be used as a prognostic target for papillary renal cell carcinoma⁷⁵. CDCA7 can promote lung adenocarcinoma proliferation by regulating the cell cycle, and its overexpression suggests a poor prognosis for triple-negative breast cancer^{76,77}. In addition, CDCA7 can promote the invasion and migration of lymphoma by regulating cytoskeletal dynamics⁷⁸. Wnt7b is a ligand of the Wnt family and is involved in the formation of many organs and tissues⁷⁹. Upregulation of Wnt7b expression levels is indicative of a poor prognosis for breast cancer⁸⁰. Although these seven key genes are closely related to the development of cancer, their biological role in kidney cancer is still not completely clear. Therefore, the roles of these genes need further study.

The results of the further stratified analysis showed that the prognostic ability of the seven-gene model was related to age and stage, and age and stage are important prognostic factors for kidney cancer that may provide reliable prognostic information and help determine the effective treatment for patients. This is consistent with recent research results^{81,82}.

Because the seven-gene model in this study could distinguish patients with a high risk of recurrence from those with low risk, it was considered that this model might be related to signaling pathways affecting kidney cancer prognosis. GSEA was performed, and the results showed that the low-expression group was mainly enriched in some biochemical-related pathways, while the tumorigenic-related pathways had higher enrichment scores in the high-expression group, which is consistent with previous results. The IL-6/Janus kinase (JAK)/signal transducer and activator of transcription 3 pathway is abnormally activated in many types of cancer and is often associated with poor clinical prognosis^{83–85}. The mediators and effectors of inflammation are important parts of a tumor's local environment. It can promote the proliferation of malignant cells and angiogenesis, interfere with the adaptive immune response, and affect the body's response to

chemotherapy⁸⁶. Interferon (IFN γ) is mainly produced by T cells and natural killer cells in response to inflammation or immune stimulation, and plasma IFN γ levels are reduced in patients with lung cancer^{87,88}. Epithelial-to-mesenchymal transition plays a vital role in tumor biology and is considered a potential therapeutic target and prognostic factor for kidney cancer⁸⁹. The main role of pancreatic beta cells is to synthesize and secrete insulin. A decrease in pancreatic beta cells is closely related to the occurrence and development of diabetes. Pancreatic islet β cells expressing LGR5 and Nanog markers may be the starting cells of pancreatic cancer^{90,91}. Protruding from the apical surface of epithelial cells are characteristic morphological features, such as microvilli. These substances promote epithelial function by expanding the surface area of the epithelium. Disturbance of these structures during their formation can lead to various diseases⁹². Bile acid metabolism is an important pathway for cholesterol catabolism⁹³. Stool bile acid levels in patients with cancer are higher than in healthy controls or those with other diseases⁹⁴. Intestinal microbes can use bile acids as messengers to regulate the antitumor immunity of patients with liver tumors⁹⁵. As the main coordinator of the immune response, IFN γ is a pleiotropic cytokine with antitumor and immunomodulatory properties⁹⁶. It plays a key role in tumor immune surveillance and stimulating antitumor immunity^{97,98}. Coagulation is closely related to the occurrence and development of gastric cancer, lung cancer, colon cancer, and ovarian cancer^{99–101}. Angiogenesis is a hallmark of cancer. Tumor growth requires blood vessels to provide nutrients and oxygen for the growth of proliferating cancer cells. Tumor blood vessels are the key targets of cancer treatments^{102,103}. This discussion not only further proves the reliability and clinical applicability of the model but also provides references for understanding the molecular mechanism of kidney cancer progression and prognosis.

To the best of our knowledge, these seven genes' potential as biomarkers has not been studied before, though studies would provide new guidance for kidney cancer prognosis. In routine clinical practice, pathological staging is a key prognostic determinant for oncologists and patients with kidney cancer. However, the different clinical outcomes of patients with kidney cancer at the same stage indicate that the current clinical staging system is insufficient for prognosis, as it is based entirely on the anatomic scope and staging system of the disease and cannot fully reflect the biological heterogeneity of patients with kidney cancer. These problems might affect the predictive accuracy of traditional systems in patients with kidney cancer.

Our findings suggest that the nomogram constructed by combining seven gene signature can clearly display the degree of risk and overall survival according to the patient's clinical stage, age, and other factors. This may be helpful for patient counseling, decision-making, and follow-up scheduling. In short, the predictive model we developed will enable patients with kidney cancer to be managed more accurately

in clinical practice. At the same time, it can help clinicians choose personalized treatment for patients.

We measured the expression of seven genes in three pairs of kidney cancer and normal tissue. Compared with normal tissue, RNASET2 and FXVD5 were highly expressed in tumor tissue, while NAT8 expression was relatively low in tumor tissue. Furthermore, there was no difference in EZH2, KLF18A, CDCA7, or WNT7B expression between tumor tissue and adjacent tissue.

In summary, existing genomic data in databases were integrated with bioinformatics technology to identify DEGs related to kidney cancer prognosis, and an mRNA prognostic model more reliable than a single-mRNA model was established. This model could well distinguish patients with high relapse risk from those with low risk, and its predictive performance was independent of age and stage. The model may not only raise new ideas for predicting the risk of kidney cancer recurrence, but it may also provide a reference for individualized treatment. However, further studies are still needed to validate the clinical applicability of the model.

Ethical Approval

This study was approved by our institutional review board.

Statement of Human and Animal Rights

This article does not contain any studies with human or animal subjects.

Statement of Informed Consent

There are no human subjects in this article and informed consent is not applicable.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Peng Wang  <https://orcid.org/0000-0002-9833-6148>

Supplemental Material

Supplemental material for this article is available online.

References

- Rini BI, Campbell SC, Escudier B. Renal cell carcinoma. *Lancet*. 2009;373(9669):1119–1132.
- Petejova N, Martinek A. Renal cell carcinoma: review of etiology, pathophysiology and risk factors. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2016;160(2):183–194.
- Abudurexiti M, Wan F, Abudurexiti M, Wang J, Zhu Y, Ye DW. Causes of death and conditional survival of renal cell carcinoma. *Front Oncol*. 2019;9:591.
- Miao D, Margolis CA, Gao W, Voss MH, Li W, Martini DJ, Norton C, Bossé D, Wankowicz SM, Cullen D, Horak C, et al. Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science*. 2018;359(6377):801–806.
- Mitchell TJ, Turajlic S, Rowan A, Nicol D, Farmery JH, O'Brien T, Martincorena I, Tarpey P, Angelopoulos N, Yates LR, Butler AP, et al. Timing the landmark events in the evolution of clear cell renal cell cancer: TRACERx renal. *Cell*. 2018;173(3):611–623.e17.
- Brussel V, Mickisch GH. Prognostic factors in renal cell and bladder cancer. *BJU Int*. 1999;83(8):902–909.
- Bonsib SM. Risk and prognosis in renal neoplasms: a pathologist's prospective. *Urol Clin North Am*. 1999;26(3):643–660.
- Medeiros LJ, Gelb AB, Weiss LM. Renal cell carcinoma. Prognostic significance of morphologic parameters in 121 cases. *Cancer*. 1988;61(8):1639–1651.
- Kim SP, Alt AL, Weight CJ, Costello BA, Cheville JC, Lohse C, Allmer C, Leibovich BC. Independent validation of the 2010 American joint committee on cancer TNM classification for renal cell carcinoma: results from a large, single institution cohort. *J Urol*. 2011;185(6):2035–2039.
- Phillips CK, Taneja SS. The role of lymphadenectomy in the surgical management of renal cell carcinoma. *Urol Oncol*. 2004;22(3):214–224.
- Golimbu M, Tessler A, Joshi P, Al-Askari S, Sperber A, Morales P. Renal cell carcinoma: survival and prognostic factors. *Urology*. 1986;27(4):291–301.
- Ljungberg B, Forsslund G, Stenling R, Zetterberg A. Prognostic significance of the DNA content in renal cell carcinoma. *J Urol*. 1986;135(2):422–426.
- Seront E, Machiels JP. Molecular biology and targeted therapies for urothelial carcinoma. *Cancer Treat Rev*. 2015;41(4):341–353.
- Sharp PA. The centrality of RNA. *Cell*. 2009;136(4):577–580.
- Midoux P, Pichon C. Lipid-based mRNA vaccine delivery systems. *Expert Rev Vaccines*. 2015;14(2):221–234.
- Weissman D. mRNA transcript therapy. *Expert Rev Vaccines*. 2015;14(2):265–281.
- Dixit D, Xie Q, Rich JN, Zhao JC. Messenger RNA methylation regulates glioblastoma tumorigenesis. *Cancer Cell*. 2017;31(4):474–475.
- Lin Q, Wang H, Lin X, Zhang W, Huang S, Zheng Y. PTPN12 affects nasopharyngeal carcinoma cell proliferation and migration through regulating EGFR. *Cancer Biother Radiopharm*. 2018;33(2):60–64.
- Shajari N, Davudian S, Kazemi T, Mansoori B, Salehi S, Khaze Shahgoli V, Shanebandi D, Mohammadi A, Duijf PH, Baradaran B. Silencing of BACH1 inhibits invasion and migration of prostate cancer cells by altering metastasis-related gene expression. *Artif Cells Nanomed Biotechnol*. 2018;46(7):1495–1504.
- Jiang C, Zhu J, Zhou P, Zhu H, Wang W, Jin Q, Li P. Overexpression of FIBCD1 is predictive of poor prognosis in gastric cancer. *Am J Clin Pathol*. 2018;149(6):474–483.

21. Masroor M, Javid J, Mir R, Prasant Y, Imtiyaz A, Mariyam Z, Mohan A, Ray PC, Saxena A. Prognostic significance of serum ERBB3 and ERBB4 mRNA in lung adenocarcinoma patients. *Tumor Biol.* 2016;37(1):857–863.
22. Qi L, Yao Y, Zhang T, Feng F, Zhou C, Xu X, Sun C. A four-mRNA model to improve the prediction of breast cancer prognosis. *Gene.* 2019;721:144100.
23. Giridhar KV, Sosa CP, Hillman DW, Sanhueza C, Dalpiaz CL, Costello BA, Quevedo FJ, Pitot HC, Dronca RS, Ertz D, Cheville JC, et al. Whole blood mRNA expression-based prognosis of metastatic renal cell carcinoma. *Int J Mol Sci.* 2017; 18(11):2326.
24. Mei J, Hu K, Peng X, Wang H, Liu C. Decreased expression of SLC16A12 mRNA predicts poor prognosis of patients with clear cell renal cell carcinoma. *Medicine.* 2019;98(30):e16624.
25. Wen L, Yu Y, Lv H, He Y, Yang B. FTO mRNA expression in the lower quartile is associated with bad prognosis in clear cell renal cell carcinoma based on TCGA data mining. *Ann Diagn Pathol.* 2019;38:1–5.
26. Ågesen TH, Sveen A, Merok MA, Lind GE, Nesbakken A, Skotheim RI, Lothe RA. ColoGuideEx: a robust gene classifier specific for stage II colorectal cancer prognosis. *Gut.* 2012; 61(11):1560–1567.
27. Karakiewicz PI, Suardi N, Capitanio U, Jeldres C, Ficarra V, Cindolo L, de La Taille A, Tostain J, Mulders PF, Bensalah K, Artibani W, et al. A preoperative prognostic model for patients treated with nephrectomy for renal cell carcinoma. *Eur Urol.* 2009;55(2):287–295.
28. Heng DY, Xie W, Regan MM, Warren MA, Golshayan AR, Sahi C, Eigl BJ, Ruether JD, Cheng T, North S, Venner P, et al. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. *J Clin Oncol.* 2009;27(34):5794–5799.
29. Gao Z, Zhang D, Duan Y, Yan L, Fan Y, Fang Z, Liu Z. A five-gene signature predicts overall survival of patients with papillary renal cell carcinoma. *PloS One.* 2019;14(3):e0211491.
30. Qiao YF, Chen CG, Yue J, Ma Z, Yu ZT. Clinical significance of preoperative and postoperative cytokeratin 19 messenger RNA level in peripheral blood of esophageal cancer patients. *Dis Esophagus.* 2016;29(8):929–936.
31. Mo R, Peng J, Xiao J, Ma J, Li W, Wang J, Ruan Y, Ma S, Hong Y, Wang C, Gao K, et al. High TXNDC5 expression predicts poor prognosis in renal cell carcinoma. *Tumor Biol.* 2016;37(7):9797–9806.
32. Almagush A, Heikkinen I, Mäkitie AA, Coletta RD, Läärä E, Leivo I, Salo T. Prognostic biomarkers for oral tongue squamous cell carcinoma: a systematic review and meta-analysis. *Br J Cancer.* 2017;117(6):856–866.
33. Breiman L. Statistical modeling: the two cultures (with comments and a rejoinder by the author). *Stat Sci.* 2001;16(3): 199–231.
34. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics Proteomics.* 2018;15(1): 41–51.
35. Wang S, Cai Y. Identification of the functional alteration signatures across different cancer types with support vector machine and feature analysis. *Biochim Biophys Acta Mol Basis Dis.* 2018;1864(6):2218–2227.
36. Hagel M, Miduturu C, Sheets M, Rubin N, Weng W, Stransky N, Bifulco N, Kim JL, Hodous B, Brooijmans N, Shutes A, et al. First selective small molecule inhibitor of FGFR4 for the treatment of hepatocellular carcinomas with an activated FGFR4 signaling pathway. *Cancer Discov.* 2015;5(4): 424–437.
37. Tan W, He J, Deng J, Yang X, Cui L, Ran R, Du G, Jiang X. Small molecule metabolite biomarkers for hepatocellular carcinoma with bile duct tumor thrombus diagnosis. *Sci Rep.* 2018;8(1):3309.
38. Lu L, Huang H, Zhou J, Ma W, Mackay S, Wang Z. BRCA1 mRNA expression modifies the effect of T cell activation score on patient survival in breast cancer. *BMC Cancer.* 2019;19(1): 387.
39. Bronte V, Brandau S, Chen SH, Colombo MP, Frey AB, Greten TF, Mandruzzato S, Murray PJ, Ochoa A, Ostrand-Rosenberg S, Rodriguez PC, et al. Recommendations for myeloid-derived suppressor cell nomenclature and characterization standards. *Nat Commun.* 2016;7(1):12150.
40. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646–674.
41. Coffelt SB, Wellenstein MD, de Visser KE. Neutrophils in cancer: neutral no more. *Nat Rev Cancer.* 2016;16(7):431–446.
42. Swierczak A, Mouchemore KA, Hamilton JA, Anderson RL. Neutrophils: important contributors to tumor progression and metastasis. *Cancer Metastasis Rev.* 2015;34(4):735–751.
43. Fridlender ZG, Sun J, Kim S, Kapoor V, Cheng G, Ling L, Worthen GS, Albelda SM. Polarization of tumor-associated neutrophil phenotype by TGF- β : “N1” versus “N2” TAN. *Cancer Cell.* 2009;16(3):183–194.
44. Antonio N, Bønnelykke-Behrndtz ML, Ward LC, Collin J, Christensen IJ, Steiniche T, Schmidt H, Feng Y, Martin P. The wound inflammatory response exacerbates growth of pre-neoplastic cells and progression to cancer. *EMBO J.* 2015;34(17): 2219–2236.
45. Spiegel A, Brooks MW, Houshyar S, Reinhardt F, Ardolino M, Fessler E, Chen MB, Krall JA, DeCock J, Zervantonakis IK, Iannello A, et al. Neutrophils suppress intraluminal NK cell-mediated tumor cell clearance and enhance extravasation of disseminated carcinoma cells. *Cancer Discov.* 2016;6(6): 630–649.
46. Welch DR, Schissel DJ, Howrey RP, Aeed PA. Tumor-elicited polymorphonuclear cells, in contrast to “normal” circulating polymorphonuclear cells, stimulate invasive and metastatic potentials of rat mammary adenocarcinoma cells. *Proc Natl Acad Sci U S A.* 1989;86(15):5859–5863.
47. Newman AC, Maddocks OD. One-carbon metabolism in cancer. *Br J Cancer.* 2017;116(12):1499–1504.
48. Pauwels AM, Trost M, Beyaert R, Hoffmann E. Patterns, receptors, and signals: regulation of phagosome maturation. *Trends Immunol.* 2017;38(6):407–422.

49. Ribas A. Adaptive immune resistance: how cancer protects from immune attack. *Cancer Discov.* 2015;5(9):915–919.
50. Spangler JB, Moraga I, Mendoza JL, Garcia KC. Insights into cytokine–receptor interactions from cytokine engineering. *Annu Rev Immunol.* 2015;33:139–167.
51. Leonard WJ, Zeng R, Spolski R. Interleukin 21: a cytokine/cytokine receptor system that has come of age. *J Leukoc Biol.* 2008;84(2):348–356.
52. Hussain M, Adah D, Tariq M, Lu Y, Zhang J, Liu J. CXCL13/CXCR5 signaling axis in cancer. *Life Sci.* 2019;227:175–186.
53. Lim SY, Yuzhalin AE, Gordon-Weeks AN, Muschel RJ. Targeting the CCL2-CCR2 signaling axis in cancer metastasis. *Oncotarget.* 2016;7(19):28697–28710.
54. Teicher BA, Fricker SP. CXCL12 (SDF-1)/CXCR4 pathway in cancer. *Clin Cancer Res.* 2010;16(11):2927–2931.
55. Waugh DJ, Wilson C. The interleukin-8 pathway in cancer. *Clin Cancer Res.* 2008;14(21):6735–6741.
56. Beauchemin N, Arabzadeh A. Carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) in cancer progression and metastasis. *Cancer Metastasis Rev.* 2013;32(3–4):643–671.
57. Fonseca I, da Luz F, Uehara IA, Silva M. Cell-adhesion molecules and their soluble forms: promising predictors of “tumor progression” and relapse in leukemia. *Tumour Biol.* 2018;40(11):1010428318811525.
58. Kobayashi H, Boelte KC, Lin PC. Endothelial cell adhesion molecules and cancer progression. *Curr Med Chem.* 2007;14(4):377–386.
59. Cooke IE, Shelling AN, Le Meuth VG, Charnock ML, Ganesan TS. Allele loss on chromosome arm 6q and fine mapping of the region at 6q27 in epithelial ovarian cancer. *Genes Chromosom Cancer.* 1996;15(4):223–233.
60. Li BC, Chan WY, Li CY, Chow C, Ng EK, Chung SC. Allelic loss of chromosome 6q in gastric carcinoma. *Diagn Mol Pathol.* 2003;12(4):193–200.
61. Negrini M, Sabbioni S, Possati L, Rattan S, Corallini A, Barbanti-Brodano G, Croce CM. Suppression of tumorigenicity of breast cancer cells by microcell-mediated chromosome transfer: studies on chromosomes 6 and 11. *Cancer Res.* 1994;54(5):1331–1336.
62. Trent JM, Stanbridge EJ, McBride HL, Meese EU, Casey G, Araujo DE, Witkowski CM, Nagle RB. Tumorigenicity in human melanoma cell lines controlled by introduction of human chromosome 6. *Science.* 1990;247(4942):568–571.
63. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. *Nature.* 2011;469(7330):343–349.
64. Xu K, Wu ZJ, Groner AC, He HH, Cai C, Lis RT, Wu X, Stack EC, Loda M, Liu T, Xu H, et al. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. *Science.* 2012;338(6113):1465–1469.
65. Li Z, Takenobu H, Setyawati AN, Akita N, Haruta M, Satoh S, Shinno Y, Chikaraishi K, Mukae K, Akter J, Sugino RP, et al. EZH2 regulates neuroblastoma cell differentiation via NTRK1 promoter epigenetic modifications. *Oncogene.* 2018;37(20):2714–2727.
66. Ma SJ, Liu YM, Zhang YL, Chen MW, Cao W. Correlations of EZH2 and SMYD3 gene polymorphisms with breast cancer susceptibility and prognosis. *Biosci Rep.* 2018;38(1):BSR20170656.
67. Wu S, Wu D, Pan Y, Liu H, Shao Z, Wang M. Correlation between EZH2 and CEP55 and lung adenocarcinoma prognosis. *Pathol Res Pract.* 2019;215(2):292–301.
68. Zhang Y, Yu X, Chen L, Zhang Z, Feng S. EZH2 overexpression is associated with poor prognosis in patients with glioma. *Oncotarget.* 2017;8(1):565–573.
69. Sun C, Zhao C, Li S, Wang J, Zhou Q, Sun J, Ding Q, Liu M, Ding G. EZH2 Expression is increased in BAP1-mutant renal clear cell carcinoma and is related to poor prognosis. *J Cancer.* 2018;9(20):3787–3796.
70. Raman P, Purwin T, Pestell R, Tozeren A. FXYD5 is a marker for poor prognosis and a potential driver for metastasis in ovarian carcinomas. *Cancer Inform.* 2015;14:113–119.
71. Besso MJ, Rosso M, Lapyckyj L, Muiola CP, Matos ML, Mercogliano MF, Schilacci R, Reventos J, Colas E, Gil A, Wernicke A, et al. FXYD5/Dysadherin, a biomarker of endometrial cancer myometrial invasion and aggressiveness: its relationship with TGF- β 1 and NF- κ B pathways. *Front Oncol.* 2019;9:1306.
72. Tassi RA, Gambino A, Ardighieri L, Bignotti E, Todeschini P, Romani C, Zanotti L, Bugatti M, Borella F, Katsaros D, Tognon G, et al. FXYD5 (Dysadherin) upregulation predicts shorter survival and reveals platinum resistance in high-grade serous ovarian cancer patients. *Br J Cancer.* 2019;121(7):584–592.
73. Liao W, Huang G, Liao Y, Yang J, Chen Q, Xiao S, Jin J, He S, Wang C. High KIF18A expression correlates with unfavorable prognosis in primary hepatocellular carcinoma. *Oncotarget.* 2014;5(21):10271–10279.
74. Luo W, Liao M, Liao Y, Chen X, Huang C, Fan J, Liao W. The role of kinesin KIF18A in the invasion and metastasis of hepatocellular carcinoma. *World J Surg Oncol.* 2018;16(1):36.
75. Feng X, Zhang M, Meng J, Wang Y, Liu Y, Liang C, Fan S. Correlating transcriptional networks to papillary renal cell carcinoma survival: a large-scale co-expression analysis and clinical validation. *Oncol Res.* 2020;28(3):285–297. doi:10.3727/096504020X15791676105394. Advance online publication.
76. Wang H, Ye L, Xing Z, Li H, Lv T, Liu H, Zhang F, Song Y. CDCA7 promotes lung adenocarcinoma proliferation via regulating the cell cycle. *Pathol Res Pract.* 2019;215(11):152559.
77. Ye L, Li F, Song Y, Yu D, Xiong Z, Li Y, Shi T, Yuan Z, Lin C, Wu X, Ren L. Overexpression of CDCA7 predicts poor prognosis and induces EZH2-mediated progression of triple-negative breast cancer. *Int J Cancer.* 2018;143(10):2602–2613.
78. Martín-Cortázar C, Chiodo Y, Jiménez RP, Bernabé M, Cayuela ML, Iglesias T, Campanero MR. CDCA7 finely tunes cytoskeleton dynamics to promote lymphoma migration and invasion. *Haematologica.* 2020;105(3):730–740.
79. Roker LA, Nemri K, Yu J. Wnt7b signaling from the ureteric bud epithelium regulates medullary capillary development. *J Am Soc Nephrol.* 2017;28(1):250–259.

80. Chen J, Liu TY, Peng HT, Wu YQ, Zhang LL, Lin XH, Lai YH. Up-regulation of Wnt7b rather than Wnt1, Wnt7a, and Wnt9a indicates poor prognosis in breast cancer. *Int J Clin Exp Pathol.* 2018;11(9):4552–4561.
81. Pal DK, Maurya AK, Jana D. Comparative study of renal cell carcinoma in patients less than 40 years of age and older age patients: a retrospective single-center study. *Indian J Cancer* 2018;55(3):297–300.
82. Williamson SR, Taneja K, Cheng L. Renal cell carcinoma staging: pitfalls, challenges, and updates. *Histopathology.* 2019;74(1):18–30.
83. Johnson DE, O’Keefe RA, Grandis JR. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nat Rev Clin Oncol.* 2018;15(4):234–248.
84. Yao Y, Ye H, Qi Z, Mo L, Yue Q, Baral A, Hoon DS, Vera JC, Heiss JD, Chen CC, Hua W, et al. B7-H4(B7x)-mediated cross-talk between glioma-initiating cells and macrophages via the IL6/JAK/STAT3 pathway lead to poor prognosis in glioma patients. *Clin Cancer Res.* 2016;22(11):2778–2790.
85. Zhao G, Zhu G, Huang Y, Zheng W, Hua J, Yang S, Zhuang J, Ye J. IL-6 mediates the signal pathway of JAK-STAT3-VEGF-C promoting growth, invasion and lymphangiogenesis in gastric cancer. *Oncol Rep.* 2016;35(3):1787–1795.
86. Candido J, Hagemann T. Cancer-related inflammation. *J Clin Immunol.* 2013;33(1):S79–S84.
87. Ni L, Lu J. Interferon gamma in cancer immunotherapy. *Cancer Med* 2018;7(9):4509–4516.
88. Wang F, Xu J, Zhu Q, Qin X, Cao Y, Lou J, Xu Y, Ke X, Li Q, Xie E, Zhang L, et al. Downregulation of IFNG in CD4(+) T cells in lung cancer through hypermethylation: a possible mechanism of tumor-induced immunosuppression. *PLoS One.* 2013;8(11):e79064.
89. Liang J, Liu Z, Zou Z, Tang Y, Zhou C, Yang J, Wei X, Lu Y. The correlation between the immune and epithelial-mesenchymal transition signatures suggests potential therapeutic targets and prognosis prediction approaches in kidney cancer. *Sci Rep.* 2018;8(1):6570.
90. Liang C, Hao F, Yao X, Qiu Y, Liu L, Wang S, Yu C, Song Z, Bao Y, Yi J, Huang Y. Hypericin maintains pdx1 expression via the erk pathway and protects islet β -cells against glucotoxicity and lipotoxicity. *Int J Biol Sci.* 2019;15(7):1472–1487.
91. Amsterdam A, Raanan C, Schreiber L, Polin N, Givol D. Lgr5 and nanog identify stem cell signature of pancreas beta cells which initiate pancreatic cancer. *Biochem Biophys Res Commun.* 2013;433(2):157–162.
92. Apodaca G. Role of polarity proteins in the generation and organization of apical surface protrusions. *Cold Spring Harb Perspect Biol.* 2018;10(1):a027813.
93. Chiang JYL, Ferrell JM. Bile acid metabolism in liver pathobiology. *Gene Expr.* 2018;18(2):71–87. doi:10.3727/105221618X15156018385515.
94. Hill MJ, Drasar BS, Williams REO, Meade TW, Cox AG, Simpson JEP, Morson BC. Faecal bile-acids and clostridia in patients with cancer of the large bowel. *Lancet.* 1975;305(7906):535–539.
95. Jia B. Commentary: gut microbiome-mediated bile acid metabolism. Regulates liver cancer via NKT cells. *Front Immunol.* 2019;10:282.
96. Castro F, Cardoso AP, Gonçalves RM, Serre K, Oliveira MJ. Interferon-gamma at the crossroads of tumor immune surveillance or evasion. *Front Immunol.* 2018;9:847.
97. Chen G, Huang AC, Zhang W, Zhang G, Wu M, Xu W, Yu Z, Yang J, Wang B, Sun H, Xia H. Exosomal PD-L1 contributes to immunosuppression and is associated with anti-PD-1 response. *Nature.* 2018;560(7718):382–386. doi:10.1038/s41586-018-0392-8.
98. Xu YP, Lv L, Liu Y, Smith MD, Li WC, Tan XM, Cheng M, Li Z, Bovino M, Aubé J, Xiong Y, et al. Tumor suppressor TET2 promotes cancer immunity and immunotherapy efficacy. *J Clin Invest.* 2019;129(10):4316–4331. Published 2019 Jul 16. doi:10.1172/JCI129317.
99. Repetto O, De Re V. Coagulation and fibrinolysis in gastric cancer. *Ann N Y Acad Sci.* 2017;1404(1):27–48. doi:10.1111/nyas.13454
100. Mitrugno A, Tassi Yunga S, Sylman JL, Zilberman-Rudenko J, Shirai T, Hebert JF, Kayton R, Zhang Y, Nan X, Shatzel JJ, Esener S, et al. The role of coagulation and platelets in colon cancer-associated thrombosis. *Am J Physiol Cell Physiol.* 2019;316(2):C264–C273. doi:10.1152/ajpcell.00367.2018.
101. Swier N, Versteeg HH. Reciprocal links between venous thromboembolism, coagulation factors and ovarian cancer progression. *Thromb Res.* 2017;150:8–18.
102. Ramjiawan RR, Griffioen AW, Duda DG. Anti-angiogenesis for cancer revisited: is there a role for combinations with immunotherapy?. *Angiogenesis.* 2017;20(2):185–204. doi:10.1007/s10456-017-9552-y.
103. Viallard C, Larrivée B. Tumor angiogenesis and vascular normalization: alternative therapeutic targets. *Angiogenesis.* 2017;20(4):409–426.