

Analysis of Genomes and Transcriptomes of Clear Cell Renal Cell Carcinomas Identifies Mutations and Gene Expression Changes in TGF-beta Pathway

Xiangyu Che

The First Affiliated Hospital of Dalian Medical University

Jianyi Li

Sun Yat-sen University

Yingkun Xu

The First Affiliated Hospital of Chongqing Medical University

Qifei Wang

The First Affiliated Hospital of Dalian Medical University

Guangzhen Wu (✉ wuguang0613@hotmail.com)

The First Affiliated Hospital of Dalian Medical University

Research Article

Keywords: TGF- β , ccRCC, TCGA, prognosis, treatment

Posted Date: April 11th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1525653/v1>

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

[Read Full License](#)

Abstract

The occurrence of clear cell renal cell carcinoma (ccRCC) is related to the change of transforming growth factor- β (TGF- β) signal pathway. We adopted an integrated approach to identify and verify the effects of changes in this pathway in ccRCC, and provide direction for finding new therapeutic targets. We performed transcriptome analysis on 539 cases of ccRCC from The Cancer Genome Atlas (TCGA) and divided the samples into different TGF- β clusters according to unsupervised hierarchical clustering. We found that 76 of the 85 TGF- β pathway genes were dysregulated, and 55 genes could be used as protective factors or risk factors to affect the prognosis of ccRCC. The survival time of tumor patients with low TGF- β score was shorter than that of tumor patients with high TGF- β score. The TGF- β score was related to the expression of ccRCC key genes and deacetylation genes. The sensitivity of tumor patients to targeted drugs was different between TGF- β high score group and TGF- β low score group. The prognostic model based on TGF- β pathway gene can predict the prognosis of ccRCC patients. Grouping patients with ccRCC according to TGF- β score is of great significance to evaluate the prognosis of patients, select targeted drugs and find new targets.

Introduction

Renal cell carcinoma(RCC) is one of the most common cancers in human beings, according to the American Cancer Society, the number of new cases of kidney cancer and the number of deaths from kidney cancer remain high in recent years^[1-3]. Clear cell renal cell carcinoma (ccRCC) accounts for 75%-80% of renal cell carcinoma^[4] and is insensitive to radiotherapy or chemotherapy^[5]. At the time of preliminary diagnosis, 20%-30% of patients with renal cell carcinoma have local or distant metastasis^[6]. For these patients, targeted drug therapy is a common treatment, and drug resistance is one of the reasons for treatment failure. A number of studies have revealed the molecular mechanisms and signal pathways of renal cell carcinoma, including TGF- β , Wnt- β -catenin and angiogenesis signal transduction. As one of the common signal pathways, TGF- β signaling pathway plays an important role in the occurrence and development of renal cell carcinoma.

It is known that TGF- β pathway plays a complex role in the occurrence and development of renal cell carcinoma and has a significant impact on tumor metastasis and prognosis^[7-8]. At present, most studies believe that TGF- β pathway is involved in the regulation of tumor as a cancer-promoting factor^[9-10]. However, some studies have shown that TGF- β can induce apoptosis of renal cancer cells, and c-Ski can weaken the anti-tumor effect of TGF- β by inhibiting TGF- β signal transduction^[11].

Here, we use the data from the The Cancer Genome Atlas (TCGA) to systematically analyze the genetic changes, prognosis and treatment-related information of TGF- β -related genes in renal cell carcinoma, and to explore the role of TGF- β signaling pathway in renal cell carcinoma.

Methods

Data acquisition and analysis

The RNA-seq transcriptgroup data for the ccRCC group were downloaded from the Genomic Data Commons (GDC) portal via R/BioConductor Package TCGAbiolink^[30], including 539 cases of clear cell renal cell carcinoma tissue and 72 cases of normal renal tissue. A total of 85 genes related to TGF pathway were obtained from KEGG TGF- β signaling pathway of GSEA website. The clinical information of cancer patients came from TCGAbiolink, including tumor size (T) status, metastatic (M) status, tumor grade, tumor stage, age and survival status. Lasso regression analysis was carried out with "Glmnet" and "Survival" packages. Univariate and multivariate Cox risk analysis of clinical features was performed with "survival" package. The correlation of immune infiltration was analyzed with "ggstatsplot" package.

Genetic alteration and survival analysis

The differential expression of TGF- β pathway genes in ccRCC and normal renal tissues was analyzed by "Limma" package, and the effect of TGF- β pathway genes on the prognosis of patients with renal cell carcinoma was analyzed by "survival" package. We downloaded the single nucleotide variations(SNV) data and expression data of TGF- β pathway genes across cancer types from TCGA database^[31] (<https://portal.gdc.cancer.gov>), analyzed them with Perl language, and visualized them with TBTools software^[32].

Cluster analysis based on TGF- β scores

Because of the great differences in gene expression profiles in the previously obtained datasets, we constructed a TGF- β Score model based on mRNA expression to show the differential expression between samples. According to the expression of mRNA in normal renal tissues, we divided the expression of mRNA in renal carcinoma tissues into three categories: TGF- β score normal group (Cluster 1), TGF- β high score group (Cluster 2) and TGF- β low score group (Cluster 3, Cluster 4). In order to further explain the relationship between gene expression levels in these three clusters, violin plots were used to describe the enrichment fraction levels in these three clusters. Kruskal-Wallis H test is used in this process. The statistical significance was set as $P < 0.05$. The "gPlots" package was used for cluster analysis by R Studio. Using R Studio's "Survival" package, we plotted survival curves for the three clusters. A heatmap was developed using "pheatmap" at R Studio to describe the relationships between the three clusters and the clinicopathological features of the patients with ccRCC. The statistical significance was set as $P < 0.05$.

Prediction of targeted drug response

We predicted the therapeutic response based on the largest publicly available pharmacogenomics database the Genomics of Drug Sensitivity in Cancer (GDSC)^[33] (<https://www.cancerrxgene.org/>). The R packet "pRRophetic" was used to realize the prediction process, the ridge regression was used to estimate the half maximum inhibitory concentration (IC50) of the sample, and the 10-fold cross-validation based on the GDSC training set was used to evaluate the prediction accuracy^[34-35]. All parameters were set by

the default values with removal of the batch effect of “combat” and tissue type of “allSolidTumours,” and duplicate gene expression was summarized as mean value^[36].

Immune cell infiltration and immunotherapy

We analyzed the immune cells quantitatively by single sample gene set enrichment (ssGSEA) analysis combined with the expression of related genes in TCGA database^[37]. The heat map used to show the correlation between the two was drawn using “ggplot2” and “dplyr” in the R language. ssGSEA analysis can apply the gene signals expressed by immune cells populations to a single sample. The 29 immune cells and regulators used in the study involved innate and acquired immunity. We used five R software packages “data.table”, “dplyr”, “tidyr”, “ggplot2” and “ggstatsplot” to analyze and plot the correlation between TGF-β score and immune substances. We selected two classical immune regulators: Type-Ⅱ-IFN response and mast cells, then used the “ggdisterstats” package to draw scatter diagram to show their correlation with TGF-β Score, respectively. Visual correlation matrix analysis was used to show the relationship between CTLA-4, PD-1 and TGF-β scores. We used tumor immune dysfunction and rejection (TIDE) algorithm (<http://tide.dfci.harvard.edu/>) and subclass mapping (<https://cloud.genepattern.org/gp>) to predict the clinical response of renal cell carcinoma to blocking immune checkpoints programmed death 1 (PD-1, PDCD-1) and cytotoxic T-lymphocyte associated protein-4 (CTLA-4)^[38-39]. Bonferroni correction is used in this process.

Establishment of Lasso regression Prognostic Model

According to the $P < 0.05$, we first selected TGF-β as the survival-related gene. Then, we use Lasso regression analysis to eliminate genes that may overfit the model. Finally, we use multivariate analysis to determine the optimal predictive factor of the model. The risk score was calculated according to the linear combination of COX coefficient and gene expression. The analysis uses the following formula: risk score = $\sum_{i=1}^N (Exp_i * Coe_i)$. N represents the number of genes, Coe_i represents the coefficient value, and Exp_i represents the level of gene expression. The median was set as the cut-off value, according to which all patients with ccRCC were divided into two groups: low-risk group and high-risk group. The overall survival time-dependent recipient operating characteristics were used to evaluate the accuracy of the prognostic model.

Compounds targeting with TGF-β pathway genes

To determine which targeted drugs may have the effect of anti-TGF-β pathway genes, we used the Baird Institute's public online tool Connectivity Map Build 02(CM)^[40] (<https://portals.broadinstitute.org/cmap/>), which allows users to predict compounds that can activate or inhibit targets based on gene expression signatures. In order to further explore the action mechanism of the compound and drug-targets, we used Connectivity Map Tools to carry out a specific analysis (<https://clue.io/>)^[41]. CMAP is a method similar to GSEA analysis, which is based on the pattern matching strategy of Kolmogorov Smirnov test to find the similarities between differentially expressed genes (DEGs). Then compare according to the DEGS ranking

list to determine the positive or negative regulatory relationship of genes, so as to generate enrichment scores (ES) from -1 to 1, and finally rank the above scores according to all the case data in the database. For each type of cancer, we could get two tables that applied the findings of the connection map to the expression characteristics of the TGF- β pathway, and then use $p < 0.05$ as our inclusion criteria to determine the average meaningful compounds for each type of tumor. The "GEOquery" package in R language was used to obtain data from the gene expression collection (GEO) database, "xlsx", "tidyverse", "plyr" and "Circlize" were used to process and analyze the data, and "pheatmap" was used to draw heat maps.

Statistical analyses

One-way ANOVA was applied to compare the expression of TGF- β pathway gene in ccRCC tissues and kidney normal tissues. Student t-test was applied to compare the expression of TGF- β pathway gene in ccRCC datasets according to age, stage, grade, T (tumor size) and M (tumor metastasis) status. The cut-off value of each risk score of carcinoma group was determined by "Survminer" package, and the patients were divided into high-risk group and low-risk group according to the best cut-off value. Rstudio software package was used for statistical analysis. $P < 0.05$ was considered to be statistically significant.

Results

The TGF- β pathway genes were significantly differentially expressed in ccRCC and were closely related to prognosis

From KEGG TGF- β signaling pathway of GSEA website, we found 85 genes related to TGF- β signaling pathway (Table S1), and analyzed the RNA sequencing data of 539 patients with human renal clear cell carcinoma and 72 normal kidney samples. We found that 76 genes of TGF- β pathway were dysregulated in ccRCC, such as TGFB1, SMAD3, SMAD4, BMP2 and so on (Figure 1A, Table S2). Furthermore, we analyzed the effects of these differentially expressed genes related to TGF- β pathway on the overall survival of patients with clear cell renal cell carcinoma. The results showed that 55 genes could significantly affect the prognosis of patients with ccRCC, of which 19 genes were risk factors for the prognosis of patients with ccRCC (risk ratio > 1) and 36 genes were protective factors (risk ratio < 1) (Figure 1B, Table S3).

Unsupervised hierarchical clustering and Prognosis analysis

The unsupervised hierarchical clustering of the data revealed four different clusters of TGF- β pathways (Figure 2A). Cluster "1" showed that the expression level of 85 genes was similar to that of normal samples, indicating that the TGF- β pathway was in a normal state. The expression level of genes related to TGF- β pathway in cluster "2" increased, indicating that TGF- β scores are high, while cluster "3" and cluster "4" showed the low expression of TGF- β pathway genes, indicating that TGF- β scores are low. Figure 2B depicts the level of TGF- β scores in four clusters and further shows the differences in TGF- β scores among them. Table S4 showed the TGF- β score of each sample. However, subsequent survival

analysis showed that the prognosis of cluster “3” was the worst and that of cluster “4” was the best (Figure 2C). Then we further classified all the samples into three groups, the “normal” group, the “KEGG-TGF- β high score group” and the “KEGG-TGF- β low score group”. That is, clusters “3” and “4” were merged into “KEGG-TGF- β low score group”. Then we analyzed the correlation between TGF- β score and clinicopathological features, and the results showed that the expression of TGF- β pathway genes was significantly correlated with fustat (Figure 2D). Further survival analysis showed that the prognosis of “KEGG-TGF- β low score group” group was the worst and the “KEGG-TGF- β high score group” group was the best (Figure 2E).

The disruption of TGF- β signaling pathway is closely related to the dysregulation of key genes and deacetylated genes in ccRCC

We explored the relationship between the expression of a variety of well-known key genes and TGF- β pathway in ccRCC (Figure 3A). VHL, TP53 and PTEN were down-regulated in the “KEGG-TGF- β low score group”. These results suggested that the destruction of TGF- β signal pathway was related to the promotion of tumor. EGFR, MYC and VEGFA and other oncogenes were highly expressed in “KEGG-TGF- β high score group”, and it may be more effective to target these genes in KEGG-TGF- β high score group. The analysis of transcriptome of ccRCC patients also showed that there was a strong correlation between the abnormal expression of sirtuin and HDAC and the abnormality of TGF- β pathway (Figure 3B).

Prediction of IC50 of different targeted drugs based on the TGF- β score

Considering that targeted drugs are commonly used in the treatment of metastatic renal cell carcinoma, we tried to evaluate the efficacy of different targeted drugs in TGF- β “KEGG-TGF- β high score group” and “KEGG-TGF- β low score group”. We have obtained a satisfactory prediction model by ridge regression on the GDSC cell line data set. Based on this model, we evaluated the IC50 of 12 targeted drugs. The results of the analysis showed that there was no significant difference in the IC50 of Pazopanib, Gefitinib, Bosutinib, and Lapatinib among the three groups (Figure 4A, 4D, 4F, 4H and 4M). Compared with KEGG-TGF- β high score group and KEGG-TGF- β low score group, the IC50 of Temsirolimus and Sunitinib in normal group was lower (Figure 4E and 4K). It is recommended that these two drugs be used in normal group. Compared with normal group and KEGG-TGF- β low score group, the IC50 of Imatinib, Nilotinib and Axitinib in KEGG-TGF- β high score group was lower (Figure 4B, 4C and 4L). These three drugs may have better results in KEGG-TGF- β high score group. Compared with normal group and KEGG-TGF- β high score group, the IC50 of Metformin, Tipifarnib and Sorafenib in KEGG-TGF- β low score group was lower (Figure 4G, 4I and 4J). These three drugs may have better results in KEGG-TGF- β low score group.

TGF- β pathway is related to immune regulation

Immune regulation plays an important role in the occurrence and development of tumors. Many genes related to TGF- β signaling pathway were associated with the infiltration of many kinds of immune cells, especially TNF, TGFB1 and IFNG genes (Figure 5A). Then we analyzed the correlation between TGF- β score and immune infiltration, and found that there was a close relationship between TGF- β score and

immune infiltration(Figure 5B), especially Type_1_IFN-response and mast cells, which were positively correlated with TGF- β score(Figure 5C and 5D). Next, we analyzed the correlation between TGF- β score and immune checkpoints, and the results showed that TGF- β score was negatively correlated with CTLA4 and PDCD1(Figure 5E). We used the TIDE algorithm to predict the response of KEGG-TGF- β low score group and KEGG-TGF- β high score group to immune checkpoint inhibitors, but after Bonferroni correction, we did not find that there was a significant difference in the response of KEGG-TGF- β low score group and KEGG-TGF- β low score group to immune checkpoint inhibitors (Figure 5F).

Construction and verification of the new TGF- β -based survival model

Firstly, TGF- β genes related to survival were screened out according to survival analysis and $P < 0.05$. Then the lasso regression model was used to analyze and determine the most reliable prognostic markers. On this basis, five genes ZFYVE9, ACVR2A, IFNG, AMH and THBS3 were selected to establish a risk characteristic model based on the minimum criterion (Figure 6A and 6B). Then, according to the median risk score, we divided patients with ccRCC into low risk group and high risk group, and studied the prognostic performance of the new survival model composed of 5 genetic risk characteristics. Kaplan-Meier survival curve analysis showed that the survival rate of patients in the high-risk group was significantly lower than that in the low-risk group (Figure 6C). In addition, we also performed ROC curve analysis to analyze the prognostic performance of the new survival model in patients with ccRCC. The AUC value of 5-year survival rate was 0.728, and that of 7-year survival rate was 0.744. The AUC value of 10-year survival rate was 0.752. It is suggested that the prognostic model of ccRCC based on TGF- β pathway gene has high predictive value (Figure 6D, 6E and 6F). In order to better understand the relationship between TGF- β pathway genes and clinicopathological characteristics of ccRCC patients, we systematically analyzed the correlation between risk scores based on five TGF- β pathway genes and clinicopathological characteristics of ccRCC patients in TCGA database. We observed that the risk score was closely related to the clinicopathological features of patients with renal clear cell carcinoma, such as T (tumor size), M (tumor metastasis), tumor grade, tumor stage and fustat(Figure 6G). Univariate Cox regression analysis showed that age, grade, stage, tumor size, tumor metastasis and risk score were associated with Overall survival in patients with ccRCC (Figure 6H, Table S5). Multivariate Cox regression analysis showed that age, grade, stage and risk score were independent risk factors that affect the prognosis of ccRCC patients (Figure 6I, Table S6). We used nomogram to predict the risk of ccRCC patients. Five-year, seven-year and ten-year survival rates can be estimated based on the patient's age, grade, stage, and risk score (Figure 6J).

TGF- β pathway genes had a wide range of genetic changes across cancer types and affected the prognosis of many cancers

In order to further explore the genetic changes of TGF- β pathway genes in pan-tumor, we analyzed the single nucleotide variation and gene expression changes of TGF- β pathway genes across cancer types. The results showed that TGF- β pathway genes had a wide range of SNV (Figure 7A, Table S7) and gene expression differences (Figure 7B, Table S8) across cancer types. Then, we analyzed the effect of TGF- β

pathway genes on the prognosis of cancer patients (Figure 7C, Table S9), and the results showed that most of TGF- β pathway genes were risk factors in other types of tumors, but they were very different in renal cell carcinoma. Most of TGF- β pathway genes were protective factors in KIRC (Kidney renal clear cell carcinoma), which was consistent with our previous analysis: the prognosis of KEGG-TGF- β low score group was worse.

Connectivity map analysis identified potential compounds / inhibitors targeting TGF- β

Considering that most TGF- β pathway genes are risk factors in tumors, we try to find compounds that can inhibit TGF- β pathway genes. We used a data-driven systematic method, Connectivity Map (CMAP), to explore the relationship between genes, compounds, and biological conditions to find candidate compounds that may target TGF- β pathway genes. We found that 54 compounds are enriched in different tumors, and they can inhibit TGF- β pathway genes (Figure 8A). At the same time, we explored the possible action mechanisms of 19 small molecular compounds, we found that the compounds involved 18 mechanisms, and two compounds had the same mechanism (Figure 8B). Therefore, we can select different compounds in different tumors and suppress TGF- β pathway genes according to different mechanisms, so as to achieve the purpose of tumor inhibition.

Discussion

Our comprehensive analysis of a large number of cases of open access renal cell carcinoma provides new insights into the key role of transforming growth factor- β pathway in the occurrence and development of renal cell carcinoma. Previous studies have reported the effects of various TGF- β pathways on renal cell carcinoma. For example, TGF- β 1 enhances the proliferation and metastatic potential of renal cell carcinoma by up-regulating lymphoid enhancer-binding factor 1/integrin α M β 2^[12], and MUC12 relies on TGF- β 1 signal to mediate the growth and invasion of renal cancer cells^[13].

However, our study found that when we analyze all the genes of TGF- β pathway as a whole, the results are surprising. Through the analysis of TCGA database, we found that 76 of the 85 TGF- β pathway genes were significantly differentially expressed between renal cell carcinoma and normal renal tissue, and 55 genes could play a pivotal role in the prognosis of patients with renal cell carcinoma. The results of this analysis make us very interested in the role of TGF- β pathway genes in renal cell carcinoma.

In many clear cell renal cell carcinoma samples, the expression of TGF- β pathway genes was not the same. We can roughly divide the renal clear cell carcinoma samples into three groups: normal group, KEGG-TGF- β high score group and KEGG-TGF- β low score group. The degree of gene expression, prognosis and response to drugs were different among the three groups. Similar to hepatocellular carcinoma^[14], the prognosis of TGF- β activation cluster was better and that of inactivation cluster was poor in renal cell carcinoma.

We have observed that there is a correlation between the expression of many well-known genes related to renal cell carcinoma and the expression of TGF- β pathway genes. VHL gene was significantly low

expressed in the KEGG-TGF- β low score group, while the loss of VHL gene function often leads to the pathogenesis of renal cell carcinoma. Similarly, well-known tumor suppressor genes such as PTEN and P53 were also low-expressed in the KEGG-TGF- β low score group, which explained the poor prognosis of the KEGG-TGF- β low score group. However, some related studies have shown that the synergistic effect of TGF- β type I receptor and hypoxia-inducible factor- α promotes the progression of renal cell carcinoma^[15], and we all know that VHL can ubiquitinate and degrade HIF- α . Obviously, there are some differences between this and our research results, which may be worthy of further study. In the KEGG-TGF- β low score group, we found some highly expressed oncogenes, such as EGFR, MYC, MTOR and VEGFA, which play a key role in the occurrence and development of renal cell carcinoma. Therefore, compared with the KEGG-TGF- β low score group, patients in the KEGG-TGF- β high score group may have a better therapeutic effect by using these oncogenes as therapeutic targets.

Acetylation and deacetylation, as common epigenetic modifications, play an important role in the occurrence and development of tumors. The analysis of TCGA database transcriptome also showed that there was a strong correlation between the aberrant expression of sirtuin and HDAC and the abnormal expression of TGF- β pathway in patients with renal cell carcinoma. SIRT1, SIRT3 and SIRT5 were significantly correlated with the high score of TGF- β , while the expression levels of SIRT6 and SIRT7 were significantly correlated with the low score of TGF- β . In the KEGG-TGF- β low score group, the expression of SIRT3 was significantly low. Previous studies have shown that the low expression of SIRT3 in renal cell carcinoma is associated with poor prognosis^[16], which fully supports our results. The transcriptome analysis of sirtuin and HDAC also indicates that the expression of sirtuin and HDAC in different TGF- β feature groups was different, so choosing different targets in different TGF- β feature groups may be more effective.

For local or distant metastatic renal cell carcinoma, targeted drug therapy is commonly used at present, and it is still a topic worth discussing which targeted drug can benefit patients the most. The vaso-rich nature of RCC has led to the approval of tyrosine kinase inhibitors targeting the VEGF signaling axis in first - and second-line therapies for metastatic RCC in the United States and the European Union, including sorafenib, sunitinib, pazopanib, and axitinib^[17-20]. However, pazopanib, also a first-line targeted agent, was similar to sunitinib in improving progression-free survival (PFS) and overall survival (OS)^[21-22]. How patients choose between these two agents is a matter of debate, as is the case with other agents. We attempted to address this issue by using TGF- β scores in patients with renal cancer. We classified the samples according to the KEGG-TGF- β score and predicted the IC50 values of a variety of targeted drugs. We observed from the results that the sensitivity of the three groups to targeted drugs was different. This suggests that choosing different targeted drugs according to the different characteristics of patients can achieve better efficacy or appropriately reduce the drug concentration so as to reduce the side effects of the drug. This has a deeper guiding significance for us, according to the different characteristics of patients for more detailed classification of patients, so that patients get more personalized treatment, which will make the treatment effect more ideal.

Immunotherapy is a hot topic in tumor therapy. TGF- β has systemic immunosuppressive effects and inhibits host immunosurveillance^[23]. Exploring the relationship between TGF- β and immunity will help us gain more tips in TGF- β targeted therapy or immunotherapy. Our results show that the expression of TGF- β pathway genes is closely related to immune cell infiltration, and the most related to TGF- β score are Type_1_IFN-response and mast cells. This discovery may be of great significance for the development of new or improved existing immunotherapy regimens. Two immune checkpoint inhibitors PD-1 inhibitor and CTLA-4 inhibitor are the main drugs approved by FDA for the treatment of advanced renal cell carcinoma^[24-25]. Recent studies have found that the combination therapy of blocking TGF-beta and PD-1/PD-L1 has achieved relatively ideal efficacy^[26-27], it has also been found that selective inhibition of TGF- β 1 produced by GARP-expressing Tregs can overcome resistance to PD-1/PD-L1 blocking in cancer^[28].

The TGF- β score was negatively correlated with the expression of CTLA4 and PDCD1, which indicated that the TGF- β low score group had higher expression of CTLA4 and PDCD1, and blocking CTLA4 and PDCD1 immune checkpoints may have a better therapeutic effect in the TGF- β low score group. However, we regret that there is no significant difference in the response of TGF- β high score group and TGF- β low score group to anti-CTLA4 and anti-PDCD1 therapy after Bonferroni correction. We do not think that the classification of TGF- β gene expression in renal cell carcinoma has no significance in guiding immunity. TGF- β gene expression is closely related to immune cell infiltration and the expression of PD-1 and CTLA4, which is worthy of further study.

We used lasso regression to establish a prognostic model of renal cell carcinoma based on TGF- β pathway gene. The results showed that the prognosis of low risk group was significantly better than that of high risk group. Multivariate COX regression analysis showed that TGF- β score was an independent risk factor for renal cell carcinoma. These results once again proved the important significance of TGF- β pathway gene in the prognosis of renal cell carcinoma.

TGF- β pathway genes also have a wide range of SNV and gene expression alterations across cancer types, and play a role in a variety of tumor types as prognostic molecules. The integration analysis of TGF- β superfamily genetic alterations in 9125 tumor samples from 33 cancer types is consistent with our study, which elucidates the salient characteristics of TGF- β -related genes in a large group of different cancer types and finds high frequency genetic alterations in the TGF- β superfamily across cancer types^[29]. TGF- β pathway genes are risk factors in most tumors, but in KIRC, most TGF- β pathway genes belong to protective factors, which makes the role of TGF- β pathway genes in KIRC different from other tumors. Therefore, the different characteristics of TGF- β pathway genes in renal cell carcinoma are worthy of our further study.

There are also some shortcomings in this study. Although TGF- β score is closely related to immune cell infiltration and immune checkpoint expression, we found that the response of TGF- β activated cluster and TGF- β inactivated cluster to immune checkpoint inhibitors was not statistically significant, which makes TGF- β score can not be used to guide immunotherapy in ccRCC patients. Moreover, our geneset does not

contain some downstream TGF- β signaling target genes, such as EMT-related genes CDH1, CDH2, Snai1 and Vim, which makes the response to pathway activity less than perfect. Because we classified the samples mainly on the basis of TGF- β score, these genes were not within the TGF- β pathway given by KEGG, so we could not give TGF- β score for these genes, so we did not use these genes.

To sum up, compared with normal renal tissues, most of the genes of TGF- β pathway are significantly differentially expressed in ccRCC, and can be used as risk factors or protective factors to affect the prognosis of patients with ccRCC. Grouping patients with renal cell carcinoma according to TGF- β score is of great significance for evaluating the prognosis of patients, selecting targeted drugs and finding new targets. People with low TGF- β score have worse prognosis, and most of the genes of TGF- β pathway are involved in the regulation of ccRCC as protective factors.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Written informed consent for publication was obtained from all participants.

Availability of data and materials

The data supporting the results of this study are available from the corresponding author upon request.

Competing interests

The authors declare that they have no conflicts of interest.

Funding

This work was supported by the Scientific Research Fund of Liaoning Provincial Education Department (No. LZ2020071), the Dalian municipal science and technology bureau 2021RQ010, and the Liaoning Province Doctoral Research Startup Fund Program (No. 2021-BS-209).

Author contributions

G Wu and X Che designed the study and analyzed the data, J Li and Q Wang analyzed the data and wrote the manuscript. Y Xu performed the data analysis and interpreted the data. All the authors checked the manuscript.

Acknowledgments

We appreciate The Cancer Genome Atlas for the open data.

References

1. Rebecca L. Siegel, Kimberly D. Miller, Jemal A. Cancer Statistics, 2018. *CA Cancer J Clin* 2018, **68**(1): 7–30.
2. Rebecca L. Siegel, Kimberly D. Miller, Jemal A. Cancer Statistics, 2019. *CA CANCER J CLIN* 2019, **69**(1): 7–34.
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA: a cancer journal for clinicians* 2020, **70**(1): 7–30.
4. Leibovich BC, Lohse CM, Crispen PL, Boorjian SA, Thompson RH, Blute ML, Cheville JC. Histological Subtype is an Independent Predictor of Outcome for Patients With Renal Cell Carcinoma. *J UROLOGY* 2010, **183**(4): 1309–1316.
5. Coppin C, Kollmannsberger C, Le L, Porzsolt F, Wilt TJ. Targeted therapy for advanced renal cell cancer (RCC): a Cochrane systematic review of published randomised trials. *BJU INT* 2011, **108**(10): 1556–1563.
6. Gong J, Maia MC, Dizman N, Govindarajan A, Pal SK. Metastasis in renal cell carcinoma: Biology and implications for therapy. *Asian Journal of Urology* 2016, **3**(4): 286–292.
7. Wang P, Chen W, Ma T, Lin Z, Liu C, Liu Y, Hou FF. lncRNA Inc-TSI Inhibits Metastasis of Clear Cell Renal Cell Carcinoma by Suppressing TGF- β -Induced Epithelial-Mesenchymal Transition. *Molecular Therapy - Nucleic Acids* 2020, **22**: 1–16.
8. Kominsky SL, Doucet M, Brady K, Weber KL. TGF- β Promotes the Establishment of Renal Cell Carcinoma Bone Metastasis. *J BONE MINER RES* 2007, **22**(1): 37–44.
9. Du G, Yan X, Chen Z, Zhang R, Tuoheti K, Bai X, Wu H, Liu T. Identification of transforming growth factor beta induced (TGFBI) as an immune-related prognostic factor in clear cell renal cell carcinoma (ccRCC). *Aging (Albany, NY.)* 2020, **12**(9): 8484–8505.
10. Bao JM, Dang Q, Lin CJ, Lo UG, Feldkoren B, Dang A, Hernandez E, Li F, Panwar V, Lee CF, Cen JJ, Guan B, Margulis V, Kapur P, Brekken RA, Luo JH, Hsieh JT, Tan WL. SPARC is a key mediator of TGF- β -induced renal cancer metastasis. *J CELL PHYSIOL* 2020.
11. Taguchi L, Miyakuni K, Morishita Y, Morikawa T, Fukayama M, Miyazono K, Ehata S. c-Ski accelerates renal cancer progression by attenuating transforming growth factor β signaling. *CANCER SCI* 2019.
12. Liu Y, Shang D. Transforming growth factor- β 1 enhances proliferative and metastatic potential by up-regulating lymphoid enhancer-binding factor 1/integrin α M β 2 in human renal cell carcinoma. *MOL CELL BIOCHEM* 2020, **465**(1–2): 165–174.
13. Gao SL, Yin R, Zhang LF, Wang SM, Chen JS, Wu XY, Yue C, Zuo L, Tang M. The oncogenic role of MUC12 in RCC progression depends on c-Jun/TGF- β signalling. *J CELL MOL MED* 2020, **24**(15): 8789–8802.
14. Chen J, Zaidi S, Rao S, Chen J, Phan L, Farci P, Su X, Shetty K, White J, Zamboni F, Wu X, Rashid A, Pattabiraman N, Mazumder R, Horvath A, Wu R, Li S, Xiao C, Deng C, Wheeler DA, Mishra B, Akbani R, Mishra L. Analysis of Genomes and Transcriptomes of Hepatocellular Carcinomas Identifies

Mutations and Gene Expression Changes in the Transforming Growth Factor- β Pathway. *GASTROENTEROLOGY* 2018, **154**(1): 195–210.

15. Mallikarjuna P, Raviprakash TS, Aripaka K, Ljungberg B, Landstrom M. Interactions between TGF-beta type I receptor and hypoxia-inducible factor-alpha mediates a synergistic crosstalk leading to poor prognosis for patients with clear cell renal cell carcinoma. *CELL CYCLE* 2019, **18**(17): 2141–2156.
16. Jeh SU, Park JJ, Lee JS, Kim DC, Do J, Lee SW, Choi SM, Hyun JS, Seo DH, Lee C, Kam SC, Chung KH, Hwa JS. Differential expression of the sirtuin family in renal cell carcinoma: Aspects of carcinogenesis and prognostic significance. *Urologic Oncology: Seminars and Original Investigations* 2017, **35**(12): 675–679.
17. Escudier B, Eisen T, Szczylik C, Oudard S, Negrier S, Chevreau C, Desai AA, Rolland F, Demkow T, Thomas E, Hutson DO, Gore M, Freeman S, Schwartz B, Shan M, Simantov R, Bukowski RM. Sorafenib in Advanced Clear-Cell Renal Cell Carcinoma. *The new England journal of medicine* 2007, **356**: 125–134.
18. Motzer RJ, Thomas E, Hutson DO, Tomczak P, Michaelson MD, Bukowski RM, Rixe O, Oudard S, Negrier S, Szczylik C, Sindy T, Kim BSIC, Bycott PW, Baum CM, Figlin RA. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *The New England Journal of Medicine* 2007, **356**: 115–124.
19. Rini BI, Escudier B, Tomczak P, Kaprin A, Szczylik C, Hutson TE, Michaelson MD, Vera A Gorbunova, Gore ME, Rusakov IG, Negrier S, Ou Y, Castellano D, Lim HY, Uemura H, Tarazi J, Cella D, Chen C, Rosbrook B, Kim S, Motzer RJ. Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial. *Lancet* 2011, **378**: 1931–1939.
20. Motzer RJ, McCann L, Deen K. Pazopanib versus Sunitinib in Renal Cancer. *The New England Journal of Medicine* 2013, **369**(20): 1970.
21. Motzer RJ, Hutson TE, Cella D, Reeves J, Hawkins R, Guo J, Nathan P, Staehler M, de Souza P, Merchan JR, Boleti E, Fife K, Jin J, Jones R, Uemura H, De Giorgi U, Harmenberg U, Wang J, Sternberg CN, Deen K, McCann L, Hackshaw MD, Crescenzo R, Pandite LN, Choueiri TK. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. *N Engl J Med* 2013, **369**(8): 722–731.
22. Motzer RJ, Hutson TE, McCann L, Deen K, Choueiri TK. Overall Survival in Renal-Cell Carcinoma with Pazopanib versus Sunitinib. *N Engl J Med* 2014, **370**(18): 1769–1770.
23. Yang L, Pang Y, Moses HL. TGF- β and immune cells: an important regulatory axis in the tumor microenvironment and progression. *TRENDS IMMUNOL* 2010, **31**(6): 220–227.
24. Xu JX, Maher VE, Zhang L, Tang S, Sridhara R, Ibrahim A, Kim G, Pazdur R. FDA Approval Summary: Nivolumab in Advanced Renal Cell Carcinoma After Anti-Angiogenic Therapy and Exploratory Predictive Biomarker Analysis. *The Oncologist* 2017, **22**(3): 311–317.
25. Rini BI, Battle D, Figlin RA, George DJ, Hammers H, Hutson T, Jonasch E, Joseph RW, McDermott DF, Motzer RJ, Pal SK, Pantuck AJ, Quinn DI, Seery V, Voss MH, Wood CG, Wood LS, Atkins MB. The society for immunotherapy of cancer consensus statement on immunotherapy for the treatment of advanced renal cell carcinoma (RCC). *J IMMUNOTHER CANCER* 2019, **7**(1).

26. Wang Y, Gao Z, Du X, Chen S, Zhang W, Wang J, Li H, He X, Cao J, Wang J. Co-inhibition of the TGF- β pathway and the PD-L1 checkpoint by pH-responsive clustered nanoparticles for pancreatic cancer microenvironment regulation and anti-tumor immunotherapy. *BIOMATER SCI-UK* 2020, **8**(18): 5121–5132.
27. Lind H, Gameiro SR, Jochems C, Donahue RN, Strauss J, Gulley JL, Palena C, Schlom J. Dual targeting of TGF- β and PD-L1 via a bifunctional anti-PD-L1/TGF- β RII agent: status of preclinical and clinical advances. *J IMMUNOTHER CANCER* 2019, **8**(1): e433.
28. de Streel G, Bertrand C, Chalon N, Liénart S, Bricard O, Lecomte S, Devreux J, Gaignage M, De Boeck G, Mariën L, Van De Walle I, van der Woning B, Saunders M, de Haard H, Vermeersch E, Maes W, Deckmyn H, Coulie PG, van Baren N, Lucas S. Selective inhibition of TGF- β 1 produced by GARP-expressing Tregs overcomes resistance to PD-1/PD-L1 blockade in cancer. *NAT COMMUN* 2020, **11**(1).
29. Korkut A, Zaidi S, Kanchi RS, Rao S, Gough NR, Schultz A, Li X, Lorenzi PL, Berger AC, Robertson G, Kwong LN, Datto M, Roszik J, Ling S, Ravikumar V, Manyam G, Rao A, Shelley S, Liu Y, Ju Z, Hansel D, de Velasco G, Pennathur A, Andersen JB, O'Rourke CJ, Ohshiro K, Jogunoori W, Nguyen B, Li S, Osmanbeyoglu HU, Ajani JA, Mani SA, Houseman A, Wiznerowicz M, Chen J, Gu S, Ma W, Zhang J, Tong P, Cherniack AD, Deng C, Resar L, Weinstein JN, Mishra L, Akbani R. A Pan-Cancer Analysis Reveals High-Frequency Genetic Alterations in Mediators of Signaling by the TGF- β Superfamily. *CELL SYST* 2018, **7**(4): 422–437.
30. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, Ceccarelli M, Bontempi G, Noushmehr H. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *NUCLEIC ACIDS RES* 2016, **44**(8): e71.
31. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemporary oncology (Poznań, Poland)* 2015, **19**(1A): A68-A77.
32. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools - an integrative toolkit developed for interactive analyses of big biological data. *Mol. MOL PLANT* 2020, **13**(8): 1194–1202.
33. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, Ramaswamy S, Futreal PA, Haber DA, Stratton MR, Benes C, McDermott U, Garnett MJ. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *NUCLEIC ACIDS RES* 2012, **41**(D1): D955-D961.
34. Geeleher P, Cox N, Huang RS, Barbour JD. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLOS ONE* 2014, **9**(9): e107468.
35. Lu X, Jiang L, Zhang L, Zhu Y, Hu W, Wang J, Ruan X, Xu Z, Meng X, Gao J, Su X, Yan F. Immune Signature-Based Subtypes of Cervical Squamous Cell Carcinoma Tightly Associated with Human Papillomavirus Type 16 Expression, Molecular Features, and Clinical Outcome. *NEOPLASIA* 2019, **21**(6): 591–601.
36. Geeleher P, Cox NJ, Huang RS. Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines. *GENOME BIOL* 2014, **15**(3): R47.

37. Zhang L, Zhao Y, Dai Y, Cheng J, Gong Z, Feng Y, Sun C, Jia Q, Zhu B. Immune Landscape of Colorectal Cancer Tumor Microenvironment from Different Primary Tumor Location. *FRONT IMMUNOL* 2018, **9**.
38. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, Liu J, Freeman GJ, Brown MA, Wucherpfennig KW, Liu XS. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *NAT MED* 2018, **24**(10): 1550–1558.
39. Hoshida Y, Brunet JP, Tamayo P, Golub TR, Mesirov JP. Subclass mapping: identifying common subtypes in independent disease data sets. *PLOS ONE* 2007, **2**(11): e1195.
40. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *SCIENCE* 2006, **313**(5795): 1929–1935.
41. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, Donahue M, Julian B, Khan M, Wadden D, Smith IC, Lam D, Liberzon A, Toder C, Bagul M, Orzechowski M, Enache OM, Piccioni F, Johnson SA, Lyons NJ, Berger AH, Shamji AF, Brooks AN, Vrcic A, Flynn C, Rosains J, Takeda DY, Hu R, Davison D, Lamb J, Ardlie K, Hogstrom L, Greenside P, Gray NS, Clemons PA, Silver S, Wu X, Zhao W, Read-Button W, Wu X, Haggarty SJ, Ronco LV, Boehm JS, Schreiber SL, Doench JG, Bittker JA, Root DE, Wong B, Golub TR. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *CELL* 2017, **171**(6): 1437–1452.

Figures

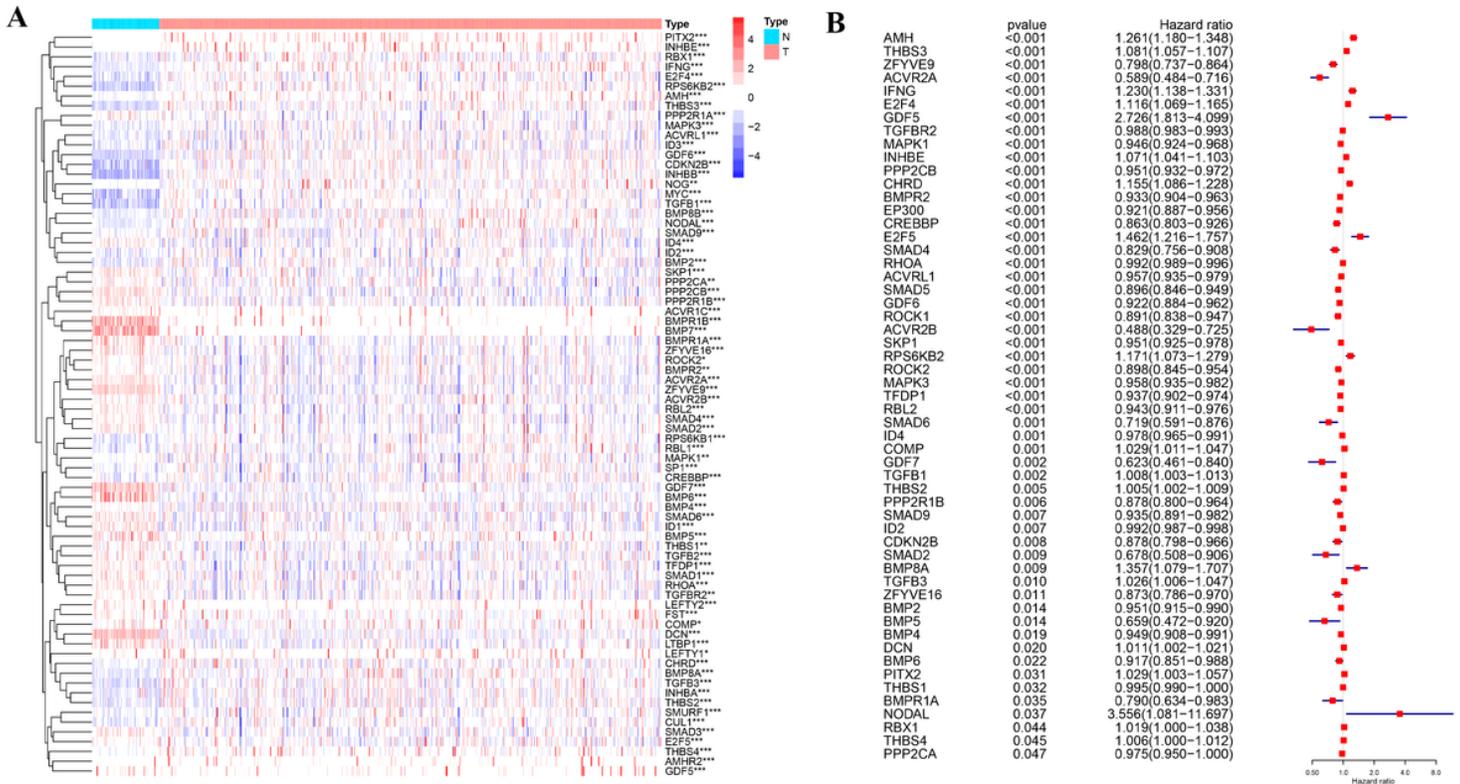


Figure 1

Differential expression of TGF- β pathway gene in ccRCC and its effect on prognosis

A Dysregulation of TGF- β pathway genes in ccRCC. **B** TGF- β pathway genes that can affect the prognosis of patients with ccRCC.

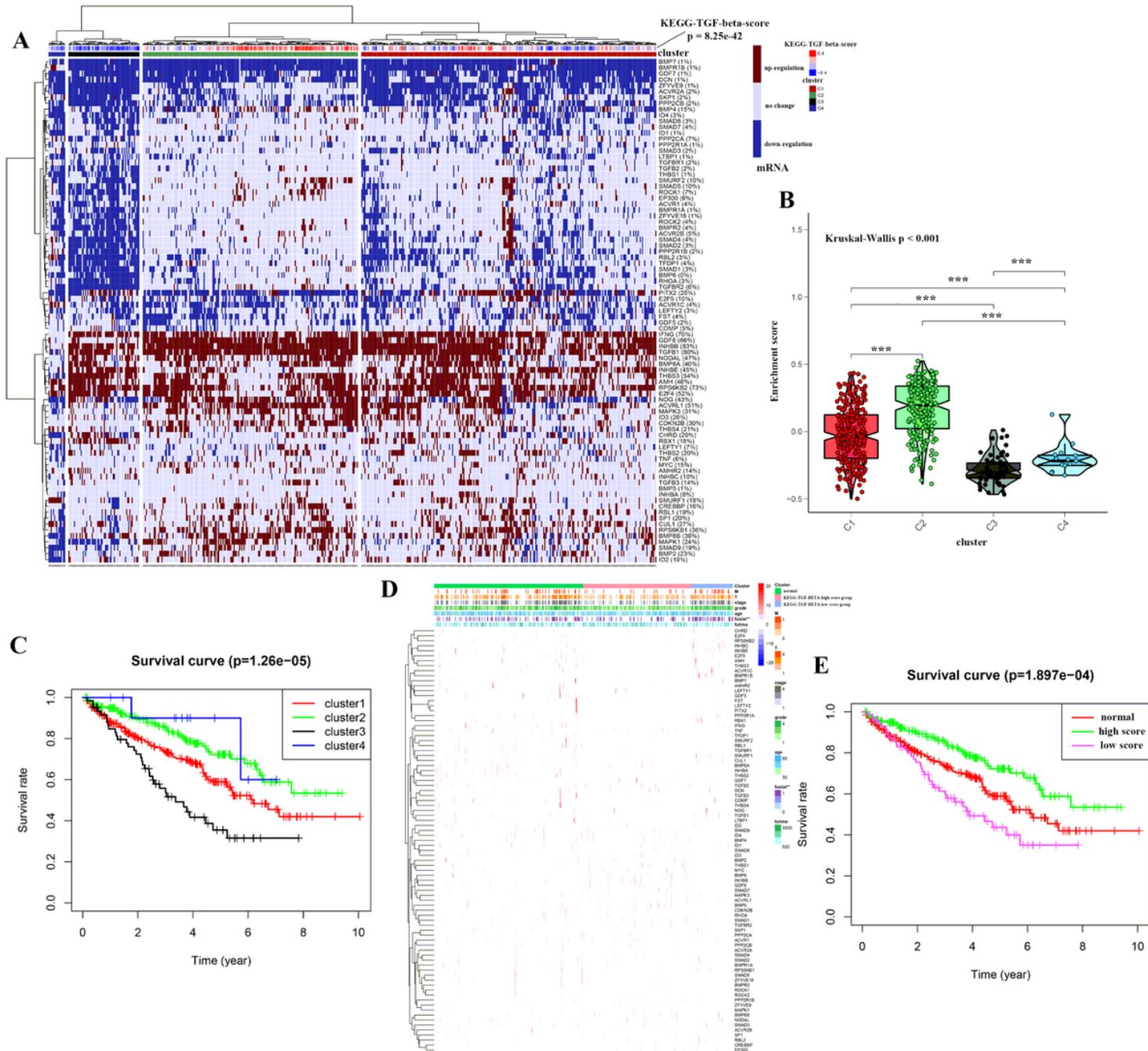


Figure 2

Unsupervised hierarchical clustering of ccRCC samples.

A Cluster analysis of transcriptome data of 539 ccRCC samples from TCGA. **B** Box plot shows the activity score of TGF- β pathway in each of the four clusters. **C** Kaplan-Meier survival curves of the four clusters. **D** The relationship between risk score and clinicopathological features. **E** Two clusters 3 and 4 were combined into one group (inactivated cluster), and compared with the activated group (activated cluster 2), there was a statistically significant difference in survival rate.

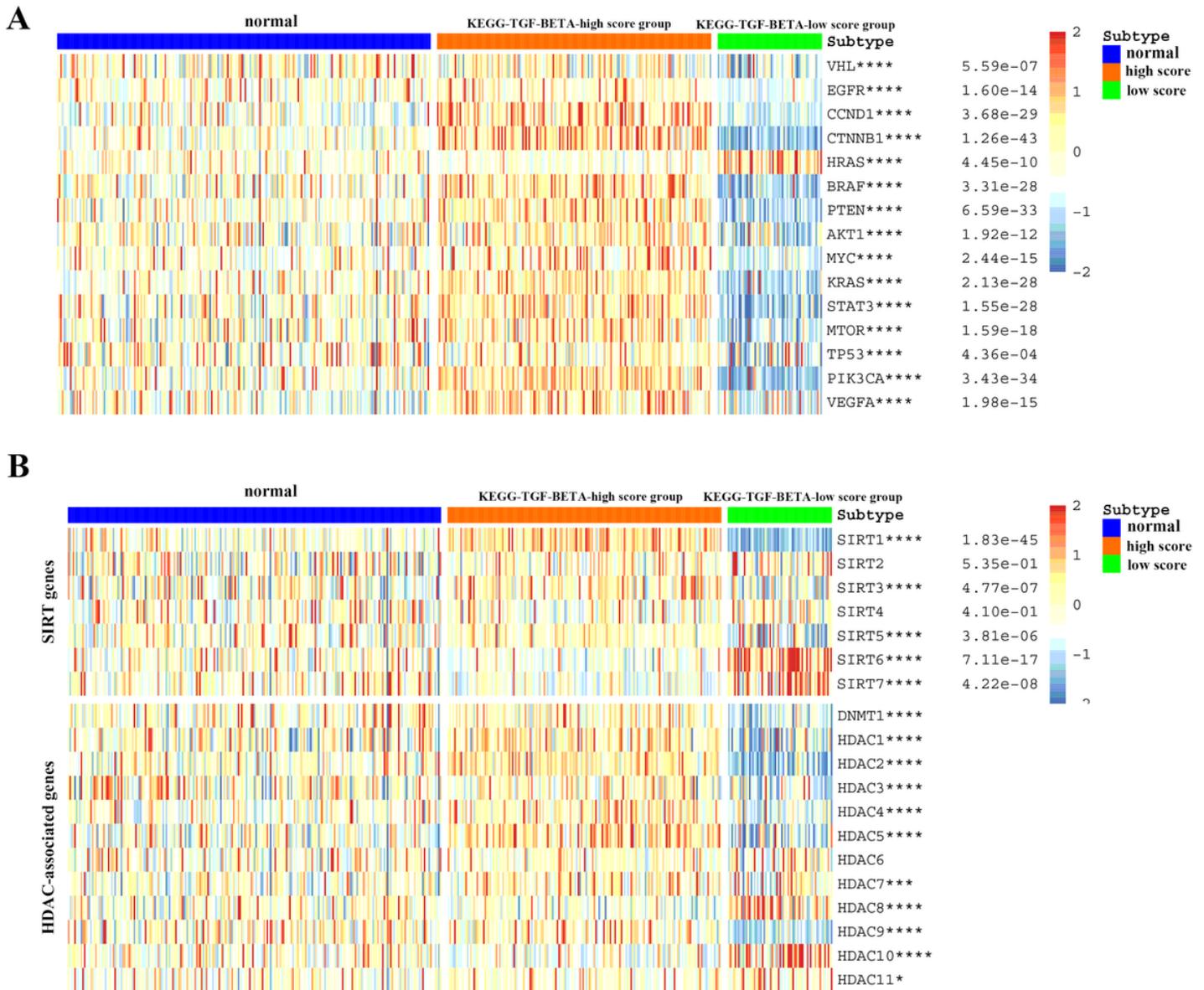


Figure 3

The relationship between the activation level of TGF- β pathway and the expression of other key genes.

A The relationship between the activation level of TGF- β pathway and the expression of oncogenes and tumor suppressor genes. **B** The relationship between the activation level of TGF- β pathway and the expression of deacetylation gene.

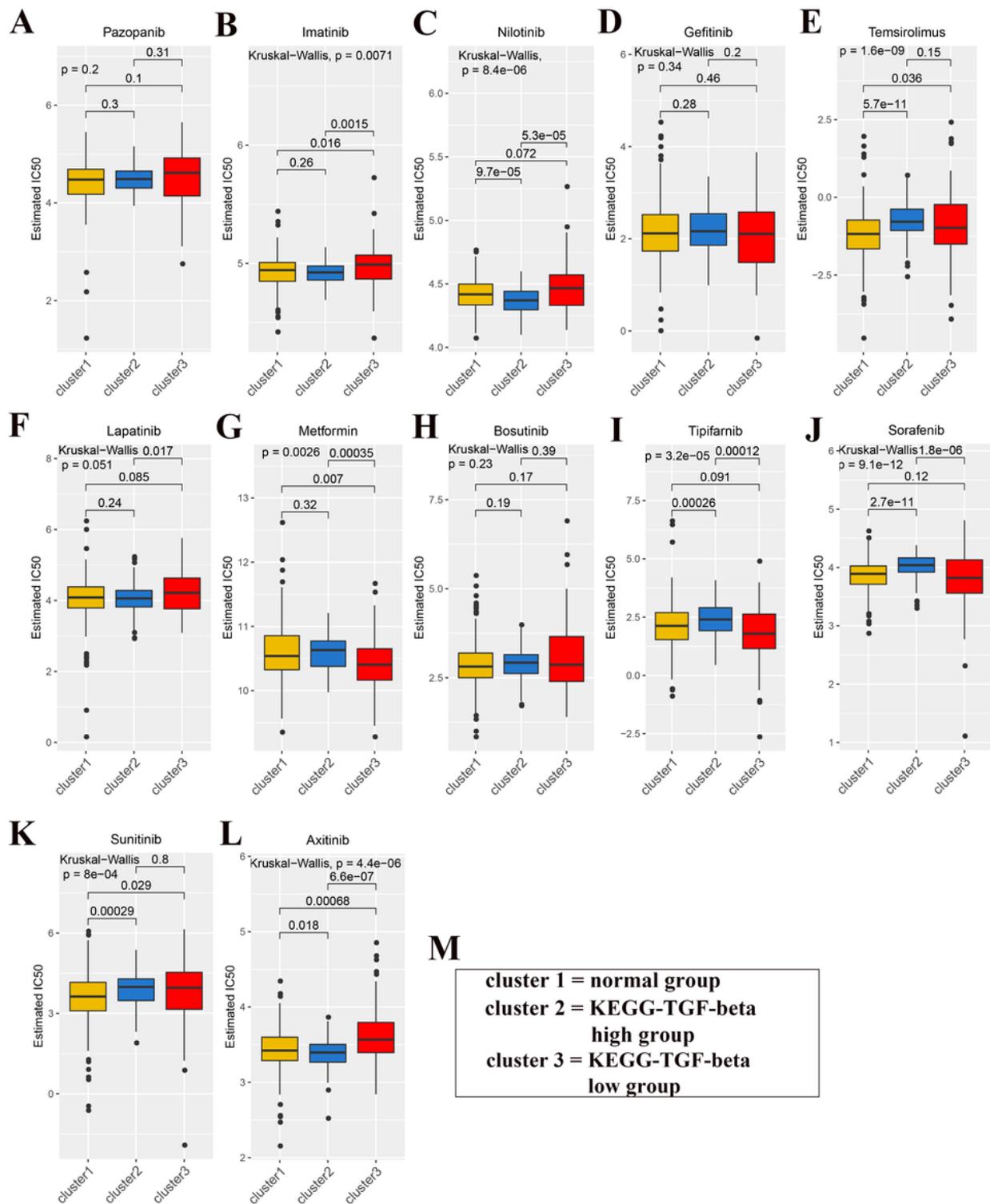


Figure 4

Prediction of IC50 value of targeted drugs.

A Pazopanib, **B** Imatinib, **C** Nilotinib, **D** Gefitinib, **E** Temsirolimus, **F** Bosutinib, **G** Metformin, **H** Lapatinib, **I** Tipifarnib, **J** Sorafenib, **K** Sunitinib, **L** Axitinib.

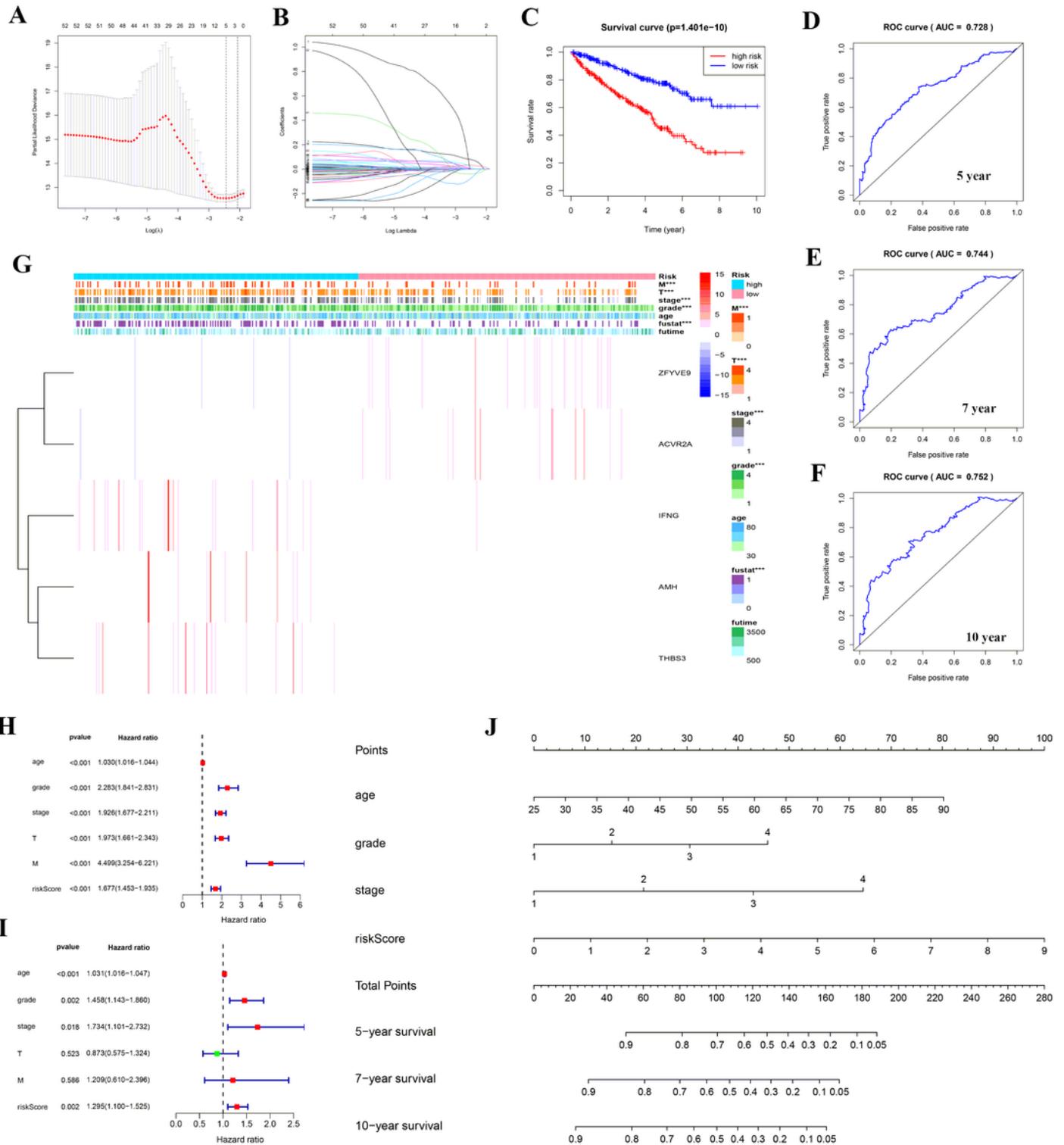


Figure 6

The establishment of prognostic model based on TGF- β in ccRCC.

A: Partial likelihood deviance was plotted against log (λ). **B:** The Lasso coefficient profiles of TGF- β in ccRCC. **C:** Grouped according to the risk scores calculated by the new survival model based on TGF- β expressions, the Kaplan-Meier survival curve showed the overall survival rate of ccRCC patients in both

the high-risk group and the low-risk group. **D-F**: ROC curve analysis showed that the new survival model was efficient in predicting prognosis, and the AUC values for 5-year survival, 7-year survival and 10-year survival were 0.728, 0.744 and 0.752, respectively. **G**: The relationship between risk score and clinicopathological features. **H**: Univariate Cox regression analysis showed that the clinicopathological parameters such as age, grade, stage, tumor size (T), tumor metastasis (M) and the risk score of the new survival model were correlated with OS in patients with renal cell carcinoma. **I**: Multivariate Cox regression analysis showed that risk score, age, grade and stage were independent risk factors affecting the prognosis of patients with clear cell renal cell carcinoma. **J** The new nomogram predicts 5-, 7- or 10-year survival in patients with ccRCC. The value of each variable has a fraction on the dot axis. The nomogram has nine lines. The second, third, fourth and fifth lines represent age, grade, stage, and risk score, respectively. The total score in the sixth row is derived from the sum of each score of age, grade, stage, and risk score, based on the total score of ccRCC patients, the five-year, seven-year and ten-year survival rate of patients can be estimated.

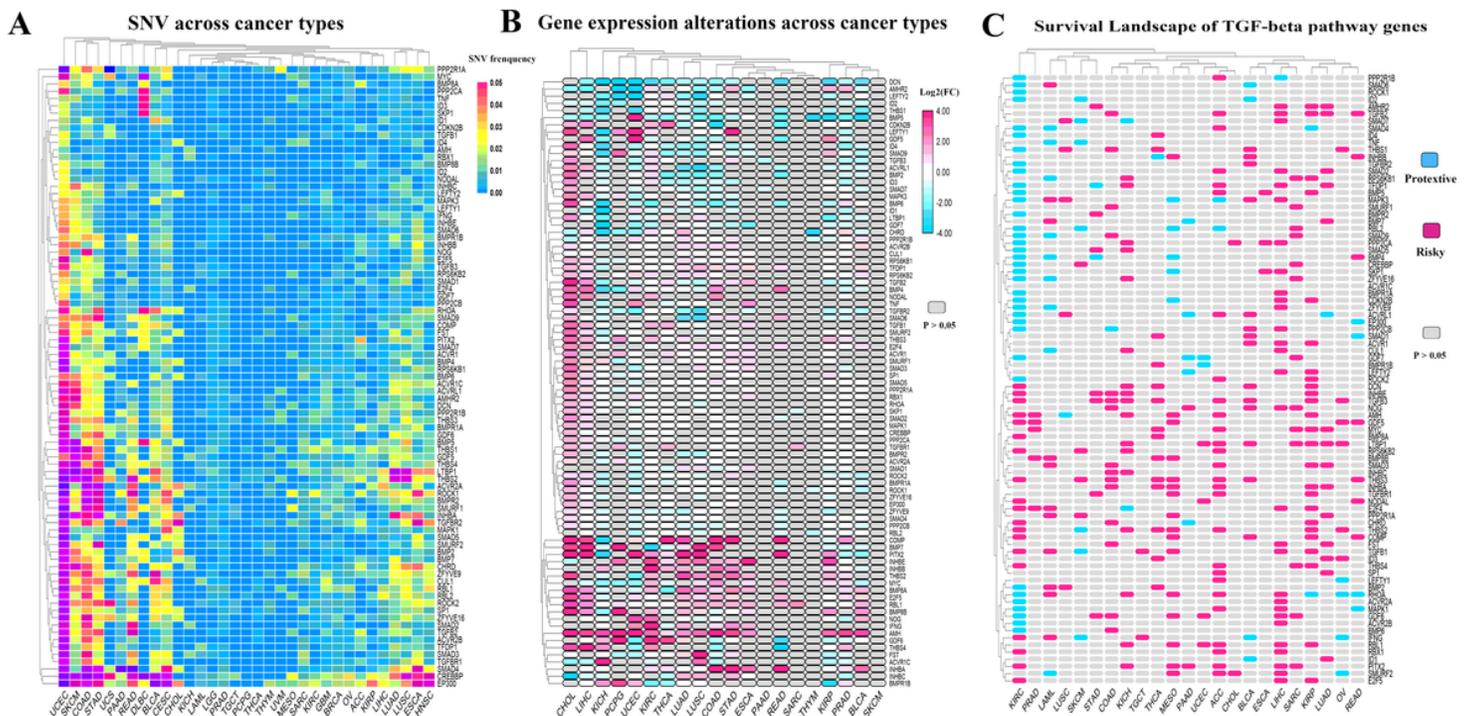


Figure 7

Genetic changes of TGF- β pathway genes across cancer types.

A SNV of TGF- β pathway genes across cancer types. **B** Alteration in expression of TGF- β pathway genes across cancer types. **C** Risk assessment of the effect of TGF- β pathway gene on prognosis.

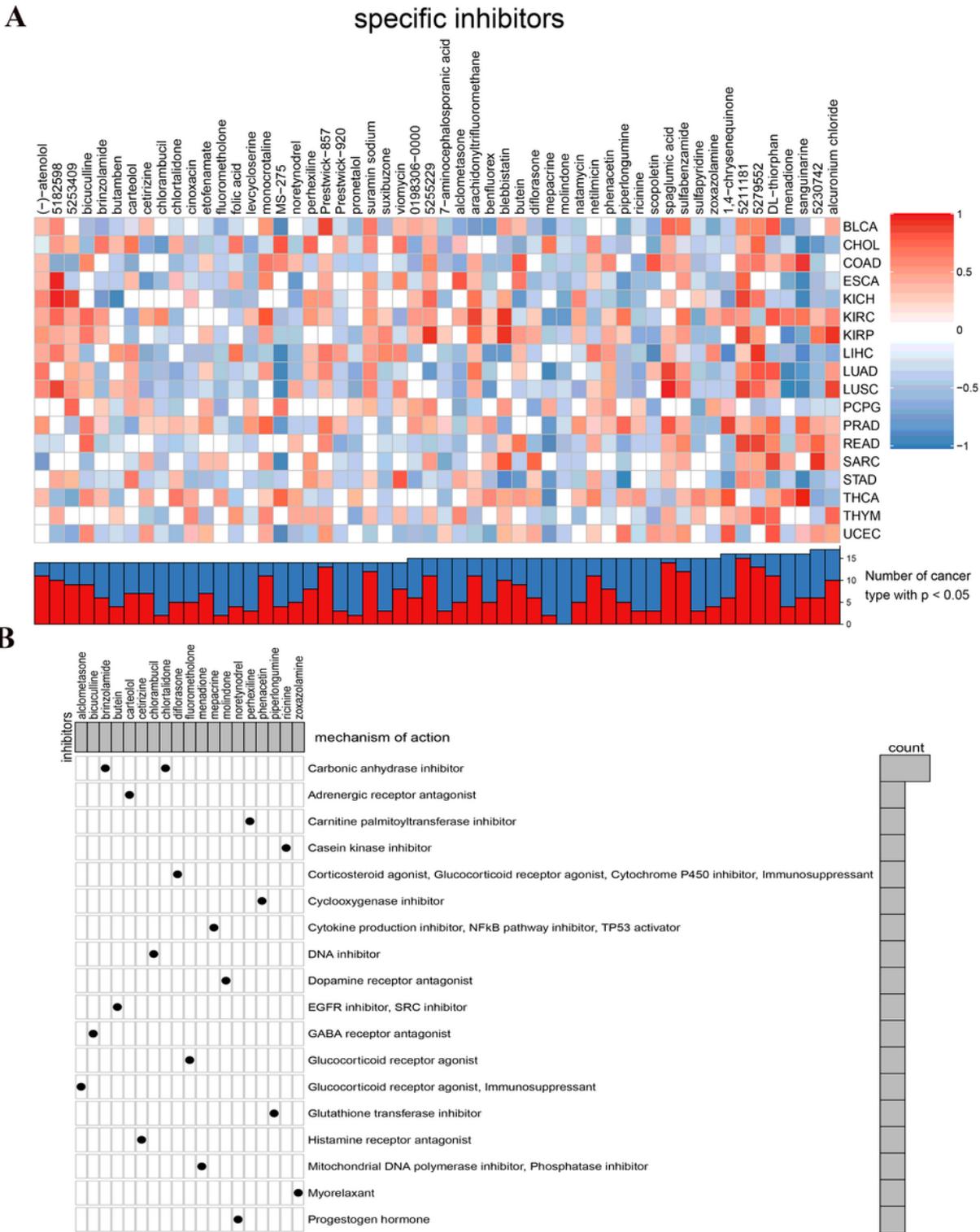


Figure 8

Connectivity map analysis identified potential compounds / inhibitors targeting TGF- β .

A Potential inhibitors of TGF- β predicted by CMAP. **B** Exploration of the mechanism of small molecular compounds.

Supplementary Files

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

- [SupplementaryMaterials.pdf](#)