

Comprehensive analysis of the expression and prognosis for COL4A family genes in kidney renal clear cell carcinoma

Jun Qian (✉ qjpaper@sina.cn)

Anhui Medical University

Bizhi Tu

Anhui Medical University

Yaya Jia

Shanxi Medical University

Laifu Wei

Anhui Medical University

Fei Gao

Anhui Medical University

Yiqun Zhang

Anhui Medical University

Qi Sha

Anhui Medical University

Research article

Keywords: COL4A, KIRC, bio-informatics, prognostic values

Posted Date: April 1st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-378296/v1>

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

[Read Full License](#)

Abstract

Background

COL4A family genes are a group of genes related to the extracellular matrix, which have been proven to be associated with various cancers. However, its relationship with kidney renal clear cell carcinoma (KIRC) has not been reported.

Methods

Hence, we obtained the data of differential expression and survival time of COL4A genes in KIRC from an online open-access database including ONCOMINE, UALCAN, GEPIA, Cancer Genome Atlas (TCGA) database, cBioPortal, Metascape, and STRING.

Results

We found significant overexpression in COL4A1, COL4A2 while COL4A 3, COL4A4, COL4A5, and COL4A6 has decreased in tumor tissues. Moreover, almost all COL4A family genes obviously correlated with individual cancer stages of KIRC; and higher expression of COL4A1/2/3/BP/4 was found to accompany better overall survival time (OS) while COL4A5 with a lower OS in KIRC patients. We also found that the COL4A genes altered group had longer OS and DFS than unaltered teams.

Introduction

Cancer is the second leading cause of death in the world. More than 1762450 new carcinoma cases are reported each year and renal cell tumors account for 2% in America [1]. However, seventy-five percent of the renal tumors were kidney renal clear cell carcinoma (KIRC), which tends to develop metastases after early radiotherapy, chemotherapy, and surgery treatment [2, 3]. Relevant studies of the KIRC pathogenesis were limited. There are sometimes variable outcomes to predict individuals' risks with KIRC using TNM stage and Fuhrman grade [4, 5]. Therefore, further research underlying molecular mechanisms of the progress and prognosis of KIRC is urgently needed to provide better clinical treatment for patients.

Collagen IV (COL4A) gene encoded a nonfibrillar protein, an essential component of the basement membrane expressing in all tissues[6, 7], which maintains organ and tissue structural integrity. The COL4A family gene was composed of 7 gene subtypes, COL4A1, COL4A2, COL4A3, COL4A3BP, COL4A4, COL4A5, and COL4A6, which encoding the different α chains of type IV collagen protein respectively [8]. Abnormal expression of collagen 4 (COL4) protein injured the renal parenchyma cell. COL4A mutations were identified as independent factors of Alport disease (AD), focal segmental glomerulosclerosis (FSGS) [9], and thin basement membrane nephropathy (TBMN) [10]. The destruction of the basement membrane provides the basis for renal inflammation and protein toxicity. New progress has been made in the study

of the pathogenesis of renal cell carcinoma, which revealed the overexpression and mislocalization of nucleoprotein SPOP produced under hypoxia's physiological condition were the core factors of the recurrence and progression in renal cancer [11]. Continuous chronic inflammatory repair reactions and hypermetabolism caused by renal tubulin hyperfiltration may further lead to increased demand for oxygen in tubular cells.

However, almost all the studies about COL4A focus on familial glomerular damage and hematuria, failing to know the correlation between the renal carcinoma and COL4A family gene. Therefore, the elucidation of KIRC from the perspective of COL4As family genes could provide a new insight for predicting clinical outcomes of patients with KIRC. Furthermore, we also analyzed the gene interaction network by obtaining 20 frequently altered similar genes to further explore the potential molecular pathway in KIRC.

Methods

ONCOMINE

ONCOMINE (<https://www.oncomine.org>) is a powerful and open access online bioinformatics database providing DNA or RNA sequences expressing analysis based on submitted tumor or normal samples [12]. In the current study, we obtained the expression data for the COL4A family gene in renal cell carcinoma from the "Expression Analysis" module of UALCAN. We set 0.05 as p-value cutoff and 1.5 as fold change to generate a p-value.

UALCAN

UALCAN (<http://ualcan.path.uab.edu/analysis.html>), a comprehensive online accessible database, providing the analysis of tumor gene expression and survival difference of prognosis based on The Cancer Genome Atlas (TCGA) and the clinical prognosis outcomes [13]. We abstracted the expression data between KIRC and normal tissues and further analyzed the survival difference in various tumor stage in patients with KIRC. Using UALCAN online analyzed resource, the cutoff of p-value was set as 0.05 in the Student's t-test.

GEPIA

GEPIA (<http://gepia.cancer-pku.cn/index.htm>), which was developed by peaking university, is a newly online analysis tool to obtain the RNA-seq expression data of more than 9,736 tumors and 8,587 normal tissue samples [14]. In our study, we analyzed differential gene expression of COL4A family among the tumor/normal tissues, different survival status, and the various pathological stage. Besides, we obtain the top 30 similar genes which closely to COL4A family genes by using a similar gene detection module. After removing the repeated genes, 178 genes were reserved to further analysis. The cutoff of p-value was set as 0.05 in Student's t-test.

Cancer Genome Atlas (TCGA) database

TCGA database (<https://portal.gdc.cancer.gov/>), a comprehensive tumor database containing gene data of 10897 samples and pathological findings, was seen as a useful online analysis tool to explore better diagnosis treatment methods by analyzing the obtained cancer genomics and clinical profiles of 34 cancer types [15].

cBioPortal

cBioPortal (<http://www.cbioportal.org>) has been used as a comprehensive online access database to explore the cancer genome data with multi-perspective [16]. Our current study obtained the genomics profiles data containing genetic mutations and putative copy-number alterations from cBioPortal based on 510 KIRC samples from the TCGA database. We set ± 1.8 . Kaplan-Meier plots as the z-score threshold of mRNA Expression (RNASeq V2 RSEM) were also applied to explore the relationship among the genetic alterations in COL4A family members and the overall survival (OS), disease-free survival (DFS) of KIRC patients. P-value as 0.05 was accepted.

Metascape

Metascape (<http://metascape.org>) is an online utility tool that provides reliable gene list annotation and enrichment analysis [17]. The 178 genes (were similar to COL4A) obtained from similar gene detection modules of Gepia were analyzed with the module of Gene Oncology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) in Metascape database. The GO module can analyze the functional roles of genes related to COL4A family members in biological processes (BP), cellular components (CC), and molecular functions (MF). And KEGG module can prescribe the pathways of the COL4A family numbers.

STRING

STRING (<http://string-db.org>) is an available protein-protein interaction (PPI) database designed for collecting, integrating, and scoring publicly available data to explore the potential protein interaction network [18]. COL4A family members and its similar gene were used to generate PPI using STRING and visualization with software Cytoscape [19]. A plug-in model Molecular Complex Detection (MCODE) of Cytoscape was used to identify the tightly connected module. We set related parameters of MCODE as: Max depth = 100, node score cut-off = 0.2, k-score = 2, and degree cut-off = 34. We considered the hub genes as the top ten genes with the highest degree of connectivity. Obviously, the higher degree connected, the more important the hub genes-maintained network structure stability.

Statistical analysis

Related statistical analysis about the correlation between mRNA expression of COL4A family genes and individuals' survival status with KIRC was conducted by SPSS (version 23.0). We remain the meaningful parameters when patients' clinicopathological parameters and mRNA expression of COL4A has a significant correlation ($p < 0.05$) for further multivariate analysis. *P*-value was accepted when less than 0.05.

Results

Expression Levels of COL4As in Patients with KIRC

According to the online analysis database Oncomine, we found differential expression of COL4As between kidney carcinoma and adjacent tissues (Fig. 1). Several reported databases confirmed that transcriptional levels of COL4A1 [20–23], COL4A2 [20, 22, 24–26], COL4A3 [20], COL3BP [22, 23] in KIRC was shown a statistically significant compare with normal tissues (Table 1). There are insufficient data of the expressed level of COL4A4, COL4A5, COL4A6 to explore the correlation between KIRC and normal tissues. However, the Ulcan and Gepia online database show an obvious difference in each COL4A gene family numbers' expression level except for COL4A3BP (Fig. 2, Fig. 3). Besides, the COL4A1, COL4A2 has increased significantly while COL4A 3, COL4A4, COL4A5, COL4A6 has decreased ($p < 0.001$ for all).

Correlation of mRNA expression levels of COL4A family genes and the tumor progression related clinicopathological parameters in KIRC patients.

The patients' individual clinicopathological parameters about cancer stages and tumor grades was explored with Gepia and Ulacn database after an over-expressed level of COL4A family members in KIRC tumors was observed. As we found in Fig. 4, there are remarkable correlations between mRNA expressions of COL4A3 ($p < 0.001$), COL4A3BP ($p < 0.001$), COL4A4 ($p < 0.001$), COL4A6 ($p = 0.0434$) and patient' pathological stages in Gepia database. However, based on the analysis of Ulcan (Fig. 5), we found a significant statistical correlation between all COL4A family genes and patients' different stages ($p < 0.01$ for all) except for COL4A3BP, which differential expression only observed in stage 4.

Prognostic value of COL4A family genes in KIRC

Using Gepia database, the prognostic value of COL4A family genes in patients were explored based on the difference of overall survival (OS) and disease-free survival (DFS) between high expression group and low expression group (Fig. 6). The OS of KIRC patients shows a statistical difference among all COL4A family genes ($p < 0.05$ for all), except for COL4A6. However, most of the COL4A family genes was shown significant correlations with kidney cancer individuals' prognosis, including COL4A3, COL4A3BP, COL4A4, COL4A5 ($p < 0.05$ for all).

The correlations between genetic mutations in COL4A family numbers and OS, DFS of KIRC patients

The online database cBioPortal was used to analyze the genetic mutations of differentially expressed COL4A family members in KIRC patients. Based on Fig. 7A, the mutation rate of COL4A1, COL4A2, COL4A3, COL4A3BP, COL4A4, COL4A5, COL4A6 genes was 8%, 9%, 9%, 11%, 7%, 9%, and 8% in 512 samples. What's more, the association between genetic mutations and the prognosis of KIRC patients was explored by Kaplan-Meier module and log-rank test in cBioPortal. And a statistically significant correlation was found between genetic mutations of COL4A family numbers and OS ($p = 0.0405$), DFS ($p = 0.0298$) in KIRC patients.

Networks Analyses and Functional Enrichment Analysis of COL4A family genes and their Neighboring Genes in KIRC patients.

After confirmed the correlation between genetic mutations in COL4A family numbers and prognosis values, the COL4A's neighbor genes (total 118) obtained from String database was used to construct PPI network to explore the interaction among neighbor genes. And the top ten interacted genes were chosen and highlighted with red by using the plug-in MCODE of Cytoscape (Fig. 8A). As shown in Fig. 8A, the neighbor genes containing COL4A1, COL4A4, COL4A6, COL4A3, COL15A1, HSPG2, COL4A5 and NID1 were the most probably involved in a different expression of COL4A family genes in KIRC patients. Based on 118 neighbor genes, the functional and pathway enrichment analyses were performed to explore the COL4A's biological classification via the online tool Metascape. The COL4A family members and their neighbor genes were significantly involved in collagen – activated tyrosine kinase receptor signaling pathway, collagen – activated signaling pathway, glomerular basement membrane development, retina vasculature development in camera – type eye and glomerulus vasculature development in biologic processes (BP); and collagen type IV trimer, basement membrane collagen trimer, network – forming collagen trimer, collagen network, and complex of collagen trimers in in cellular components (CC); and extracellular matrix structural constituent conferring tensile strength, GABA receptor binding, myosin binding, extracellular matrix structural constituent, and growth factor binding in molecular function (MF); and collagen – activated tyrosine kinase receptor signaling pathway, Anchoring fibril formation, Crosslinking of collagen fibrils, collagen – activated signaling pathway, and glomerular basement membrane development in KEGG pathway.

Discussion

Previous researches on COL4A family genes mostly focused on renal parenchyma changes other than the tumor. However, in the current study, the significant correlation between COL4A family genes and renal cell carcinoma was observed (Fig. 1). And COL4A genes play an important role in the occurrence and development of KIRC (Fig. 2–5). It is reported that COL4A increases angiogenesis in tumor tissue through the interferon genes (STING) signaling pathway [27]. What's more, the COL4A carry mutations in multiple forms involving the immune cell infiltration, which expression levels have a significant negative correlation to the patients' survival time with cervical cell carcinoma [28]. Besides, high methylation levels of COL4A family genes are also found in colorectal cancer [29]. Since the alteration of COL4A has been confirmed in tumor tissues, more attention should be paid to COL4A and its' related genes to further explain the underlying mechanism in KIRC.

To further explore the correlation between COL4As family genes expressing and KIRC's progression and prognosis, we analyzed the distinct COL4As family members and their correlation to clinical parameters in KIRC patients. Until now, there are no reported researches about the role of COL4A family genes in the KIRC. Based on Fig. 2 and Fig. 3, we found that 6 out of 7 genes were differentially expressed in KIRC (downregulation of COL4A1 and COL4A2, upregulation of COL4A3, COL4A4, COL4A5, COL4A6). Moreover, the expression of COL4A1, COL4A2, COL4A3, COL4A4, COL4A5, and COL4A6 was found increasing when the tumor progressed (Fig. 5). We found that the better prognosis (better overall survival time) was accompanied by low expression of COL4A5 and high expression of COL4A1, COL4A2, COL4A3 and COL4A4 (Fig. 6). As for disease-free survival (DFS), the better results were observed in the low

expression of COL4A5 and high expression of COL4A3 and COL4A4 (Fig. 6). Based on those observations, the COL4A5 might act as a protective factor in the prognosis of KIRC patients, while COL4A3 and COL4A4 could be treated as a risk factor.

Significant mutation of COL4A genes in KIRC was found in Fig. 7. The mutation of COL4A genes was confirmed as an independent indicator in the poor prognosis of renal nonneoplastic diseases, such as FSGS[9], Alport syndrome [30], and familial hematuria [31]. Moreover, the tumorigenesis and progress of renal cell carcinoma involved in gene-altered [32]. Although the laboratory information of mutation about COL4A family members in KIRC is limited, the more altered COL4A genes result in shorter OS and DFS time in patients (Fig. 7B, 7C). We found different degrees of correlation among the differentially expressed COL4A genes, indicating that these altered genes may have an adverse role in the tumorigenesis of renal tumors.

The GO enrichment analysis and KEGG pathway enrichment analysis shows that COL4A family members and its' related 118 similar genes were primarily involved in collagen – activated tyrosine kinase receptor signaling pathway and affecting extracellular matrix structural constituent (Fig. 8). Overexpression of collagen fiber deposited in the stroma plays an unfavorable role in tumor prognosis of ductal carcinoma in situ (DCIS) [33]. Cumulating evidence has confirmed a close connection between collagen fiber component and myxoid stroma, which has different expression in several tumors [34–37]. Those reported data may provide a promising target on potential drug therapeutic methods by regulating collagen fiber expression in KIRC.

In the molecular interaction network (Fig. 8), we also found that heparan sulfate proteoglycan 2 (HSPG2) gene is closely involved in the process of renal cell carcinoma. Notably, similar to COL4A genes, the HSPG2 also widely expressed in all basement membranes including those epithelial and endothelial cells to protect the integrity of the extracellular matrix [38–40]. For NID1, also an extracellular matrix protein that regulated the activation of NK cells has been found excessively expression in basal cell carcinoma (BCC) [41], breast cancer [42], and lung cancer [43]. Reported studies have already indicated that NID1 was an indicator of prognosis in several tumors [42]. The downregulated NID1 could improve cancer patients' survival outcomes by deactivating the MET, which may prohibit the migration and invasion of cancer cells [44]. As shown in Fig. 8, eight extracellular matrix-associated genes COL4A1, COL4A4, COL4A6, COL4A3, COL15A1, HSPG2, COL4A5, and NID1 together closely involving in KIRC. These genes may reveal an extraordinary correlation between the extracellular matrix and renal cell carcinoma. These gene-related proteins may indicate a potential mechanism in KIRC and provide a new sight in further research.

Limitations

Some limitations need to be recognized in the current study. Firstly, Our data were obtained from the online public database, and its analysis module may have a decisive impacts on our analysis progress. So our results should be confirmed by further studies. Secondly, we did not confirm the potential

diagnostic and therapeutic values of COL4A family genes in KIRC patients because there are no lab results.

Conclusion

The COL4A family genes play an important role in KIRC. Low expression of COL4A5 and high expression of COL4A3 and COL4A4 may be a positive indicator for patients with KIRC. The high mutation of COL4A foreshadowing a poor prognosis. The relationship between extracellular matrix-related genes and KIRC was worthy of further studies.

Declarations

Ethics approval and consent to participate:

Not applicable.

Consent to publish:

Not applicable.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Competing interests

Author(s) declare(s) that there is no conflict of interest

Funding

This project was supported by the National Natural Science Foundation of China (grant numbers 81471273 and 81671204), the Foundation of Supporting Program for the Excellent Young Faculties in the University of Anhui Province in China. Grants for Scientific Research of BSKY from the First Affiliated Hospital of Anhui Medical University; and Grants for Outstanding Youth from the First Affiliated Hospital of Anhui Medical University.

Author contribution

JQ conceived the study idea. BZ collected the data. BZ, YY, and LF contributed to the analysis of the data as well as wrote the initial draft with all authors providing critical feedback and edits to subsequent revisions. All authors approved the final draft of the manuscript. All authors are accountable for all aspects of the work in ensuring related questions accuracy or integrity. Any parts of the work are

appropriately investigated and resolved. JQ is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Acknowledgments

We thank all the people who offer help for this study. And thank the Department of Spine Surgery, The First Affiliated Hospital of Anhui Medical University, for its grateful supports.

References

1. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2019**. *CA Cancer J Clin* 2019, **69**(1):7-34.
2. Rini BI, Campbell SC, Escudier B: **Renal cell carcinoma**. *The Lancet* 2009, **373**(9669):1119-1132.
3. Hsieh JJ, Purdue MP, Signoretti S, Swanton C, Albiges L, Schmidinger M, Heng DY, Larkin J, Ficarra V: **Renal cell carcinoma**. *Nature reviews Disease primers* 2017, **3**:17009.
4. Ficarra V, Novara G, Galfano A, Brunelli M, Cavalleri S, Martignoni G, Artibani W: **The 'Stage, Size, Grade and Necrosis' score is more accurate than the University of California Los Angeles Integrated Staging System for predicting cancer-specific survival in patients with clear cell renal cell carcinoma**. *Bju Int* 2009, **103**(2):165-170.
5. Zigeuner R, Hutterer G, Chromecki T, Imamovic A, Kampel-Kettner K, Rehak P, Langner C, Pummer K: **External Validation of the Mayo Clinic Stage, Size, Grade, and Necrosis (SSIGN) Score for Clear-Cell Renal Cell Carcinoma in a Single European Centre Applying Routine Pathology**. *European urology* 2010, **57**(1):102-109.
6. Durrani-Kolarik S, Manickam K, Chen B: **COL4A1 Mutation in a Neonate With Intrauterine Stroke and Anterior Segment Dysgenesis**. *Pediatr Neurol* 2017, **66**:100-103.
7. Vahedi K, Alamowitch S: **Clinical spectrum of type IV collagen (COL4A1) mutations: a novel genetic multisystem disease**. *Curr Opin Neurol* 2011, **24**(1):63-68.
8. Savige J, Ariani F, Mari F, Bruttini M, Renieri A, Gross O, Deltas C, Flinter F, Ding J, Gale DP *et al*: **Expert consensus guidelines for the genetic diagnosis of Alport syndrome**. *Pediatr Nephrol* 2019, **34**(7):1175-1189.
9. Yao T, Udwan K, John R, Rana A, Haghighi A, Xu L, Hack S, Reich HN, Hladunewich MA, Cattran DC *et al*: **Integration of Genetic Testing and Pathology for the Diagnosis of Adults with FSGS**. *Clin J Am Soc Nephrol* 2019, **14**(2):213-223.
10. Gast C, Pengelly RJ, Lyon M, Bunyan DJ, Seaby EG, Graham N, Venkat-Raman G, Ennis S: **Collagen (COL4A) mutations are the most frequent mutations underlying adult focal segmental glomerulosclerosis**. *Nephrology Dialysis Transplantation* 2016, **31**(6):961-970.
11. Guo ZQ, Zheng T, Chen BE, Luo C, Ouyang SS, Gong SZ, Li JF, Mao LL, Lian FL, Yang Y *et al*: **Small-Molecule Targeting of E3 Ligase Adaptor SPOP in Kidney Cancer**. *Cancer Cell* 2016, **30**(3):474-484.
12. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, Chinnaiyan AM: **ONCOMINE: a cancer microarray database and integrated data-mining platform**. *Neoplasia (New*

York, NY) 2004, **6**(1):1-6.

13. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, Varambally S: **UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses**. *Neoplasia (New York, NY)* 2017, **19**(8):649-658.
14. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z: **GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses**. *Nucleic Acids Res* 2017, **45**(W1):W98-W102.
15. Tomczak K, Czerwinska P, Wiznerowicz M: **The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge**. *Contemp Oncol (Pozn)* 2015, **19**(1A):A68-77.
16. Gao JJ, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun YC, Jacobsen A, Sinha R, Larsson E *et al*: **Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal**. *Science signaling* 2013, **6**(269).
17. Zhou YY, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK: **Metascape provides a biologist-oriented resource for the analysis of systems-level datasets**. *Nat Commun* 2019, **10**.
18. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P *et al*: **STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets**. *Nucleic Acids Research* 2019, **47**(D1):D607-D613.
19. Su G, Morris JH, Demchak B, Bader GD: **Biological network exploration with Cytoscape 3**. *Curr Protoc Bioinformatics* 2014, **47**:8 13 11-24.
20. Jones J, Otu H, Spentzos D, Kolia S, Inan M, Beecken WD, Fellbaum C, Gu XS, Joseph M, Pantuck AJ *et al*: **Gene signatures of progression and metastasis in renal cell cancer**. *Clinical Cancer Research* 2005, **11**(16):5730-5739.
21. Beroukhim R, Brunet JP, Di Napoli A, Mertz KD, Seeley A, Pires MM, Linhart D, Worrell RA, Moch H, Rubin MA *et al*: **Patterns of gene expression and copy-number alterations in von-hippel lindau disease-associated and sporadic clear cell carcinoma of the kidney**. *Cancer Res* 2009, **69**(11):4674-4681.
22. Lenburg ME, Liou LS, Gerry NP, Frampton GM, Cohen HT, Christman MF: **Previously unidentified changes in renal cell carcinoma gene expression identified by parametric analysis of microarray data**. *BMC Cancer* 2003, **3**:31.
23. Gumz ML, Zou H, Kreinest PA, Childs AC, Belmonte LS, LeGrand SN, Wu KJ, Luxon BA, Sinha M, Parker AS *et al*: **Secreted frizzled-related protein 1 loss contributes to tumor phenotype of clear cell renal cell carcinoma**. *Clinical Cancer Research* 2007, **13**(16):4740-4749.
24. Gumz ML, Zou H, Kreinest PA, Childs AC, Belmonte LS, LeGrand SN, Wu KJ, Luxon BA, Sinha M, Parker AS *et al*: **Secreted frizzled-related protein 1 loss contributes to tumor phenotype of clear cell renal cell carcinoma**. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2007, **13**(16):4740-4749.

25. Beroukhim R, Brunet JP, Di Napoli A, Mertz KD, Seeley A, Pires MM, Linhart D, Worrell RA, Moch H, Rubin MA *et al*: **Patterns of Gene Expression and Copy-Number Alterations in von-Hippel Lindau Disease-Associated and Sporadic Clear Cell Carcinoma of the Kidney.** *Cancer Research* 2009, **69**(11):4674-4681.
26. Higgins JPT, Shinghal R, Gill H, Reese JH, Terris M, Cohen RJ, Fero M, Pollack JR, van de Rijn M, Brooks JD: **Gene Expression Patterns in Renal Cell Carcinoma Assessed by Complementary DNA Microarray.** *The American journal of pathology* 2003, **162**(3):925-932.
27. Yang H, Lee WS, Kong SJ, Kim CG, Kim JH, Chang SK, Kim S, Kim G, Chon HJ, Kim C: **STING activation reprograms tumor vasculatures and synergizes with VEGFR2 blockade.** *J Clin Invest* 2019, **129**(10):4350-4364.
28. Zhou CK, Li CM, Peng SQ, Zhou LC, Li H: **Comprehensive Analysis of the Relationships Between Tumor Mutation Burden With Immune Infiltrates in Cervical Cell Carcinoma.** *Front Mol Biosci* 2020, **7**.
29. Naumov VA, Generozov EV, Zaharjevskaya NB, Matushkina DS, Larin AK, Chernyshov SV, Alekseev MV, Shelygin YA, Govorun VM: **Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips.** *Epigenetics-Us* 2013, **8**(9):921-934.
30. Savige J, Ariani F, Mari F, Bruttini M, Renieri A, Gross O, Deltas C, Flinter F, Ding J, Gale DP *et al*: **Expert consensus guidelines for the genetic diagnosis of Alport syndrome.** *Pediatr Nephrol* 2019, **34**(7):1175-1189.
31. Li Y, Wang Y, He Q, Dang X, Cao Y, Wu X, Mo S, He X, Yi Z: **Genetic mutational testing of Chinese children with familial hematuria with biopsyproven FSGS.** *Mol Med Rep* 2018, **17**(1):1513-1526.
32. Roy I, Getschman AE, Volkman BF, Dwinell MB: **Exploiting agonist biased signaling of chemokines to target cancer.** *Mol Carcinog* 2017, **56**(3):804-813.
33. Conklin MW, Gangnon RE, Sprague BL, Van Gemert L, Hampton JM, Eliceiri KW, Bredfeldt JS, Liu YM, Surachaicharn N, Newcomb PA *et al*: **Collagen Alignment as a Predictor of Recurrence after Ductal Carcinoma In Situ.** *Cancer Epidem Biomar* 2018, **27**(2):138-145.
34. Sutton BJ, Laudadio J: **Aggressive angiomyxoma.** *Archives of pathology & laboratory medicine* 2012, **136**(2):217-221.
35. Santiago T, Clay MR, Allen SJ, Orr BA: **Recurrent BCOR internal tandem duplication and BCOR or BCL6 expression distinguish primitive myxoid mesenchymal tumor of infancy from congenital infantile fibrosarcoma.** *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2017, **30**(6):884-891.
36. Hwang S, Kim YD, Siapno DL, Suh YL, Woo KI: **Fibromyxoma of the Orbit.** *The Journal of craniofacial surgery* 2019, **30**(1):e19-e21.
37. Howitt BE, Fletcher CD: **Mammary-type Myofibroblastoma: Clinicopathologic Characterization in a Series of 143 Cases.** *The American journal of surgical pathology* 2016, **40**(3):361-367.
38. Nicole S, Davoine CS, Topaloglu H, Cattolico L, Barral D, Beighton P, Hamida CB, Hammouda H, Cruaud C, White PS *et al*: **Perlecan, the major proteoglycan of basement membranes, is altered in**

patients with Schwartz-Jampel syndrome (chondrodystrophic myotonia). *Nature genetics* 2000, **26**(4):480-483.

39. Arikawa-Hirasawa E, Wilcox WR, Le AH, Silverman N, Govindraj P, Hassell JR, Yamada Y: **Dyssegmental dysplasia, Silverman-Handmaker type, is caused by functional null mutations of the perlecan gene.** *Nature genetics* 2001, **27**(4):431-434.
40. Martinez JR, Dhawan A, Farach-Carson MC: **Modular Proteoglycan Perlecan/HSPG2: Mutations, Phenotypes, and Functions.** *Genes (Basel)* 2018, **9**(11).
41. Hirakawa Y, Futaki S, Tanizaki H, Furukawa F, Maemura K, Kondo Y, Moriwaki S: **Enhanced expression of nidogen 1 around the nest of basal cell carcinoma compared with that around squamous cell carcinoma.** *Med Mol Morphol* 2019, **52**(2):99-105.
42. Urooj T, Wasim B, Mushtaq S, Haider G, Shah SNN, Ghani R, Qureshi MFH: **Increased NID1 Expression among Breast Cancer Lung Metastatic Women; A Comparative Analysis between Naive and Treated Cases.** *Recent Pat Anti-Canc* 2020, **15**(1):59-69.
43. Aleckovic M, Wei Y, Leroy G, Sidoli S, Liu DD, Garcia BA, Kang YB: **Identification of Nidogen 1 as a lung metastasis protein through secretome analysis.** *Gene Dev* 2017, **31**(14):1439-1455.
44. Zhou SQ, Zhang S, Wang L, Huang SL, Yuan Y, Yang J, Wang H, Li XH, Wang P, Zhou L *et al*: **BET protein inhibitor JQ1 downregulates chromatin accessibility and suppresses metastasis of gastric cancer via inactivating RUNX2/NID1 signaling.** *Oncogenesis* 2020, **9**(3).

Figures

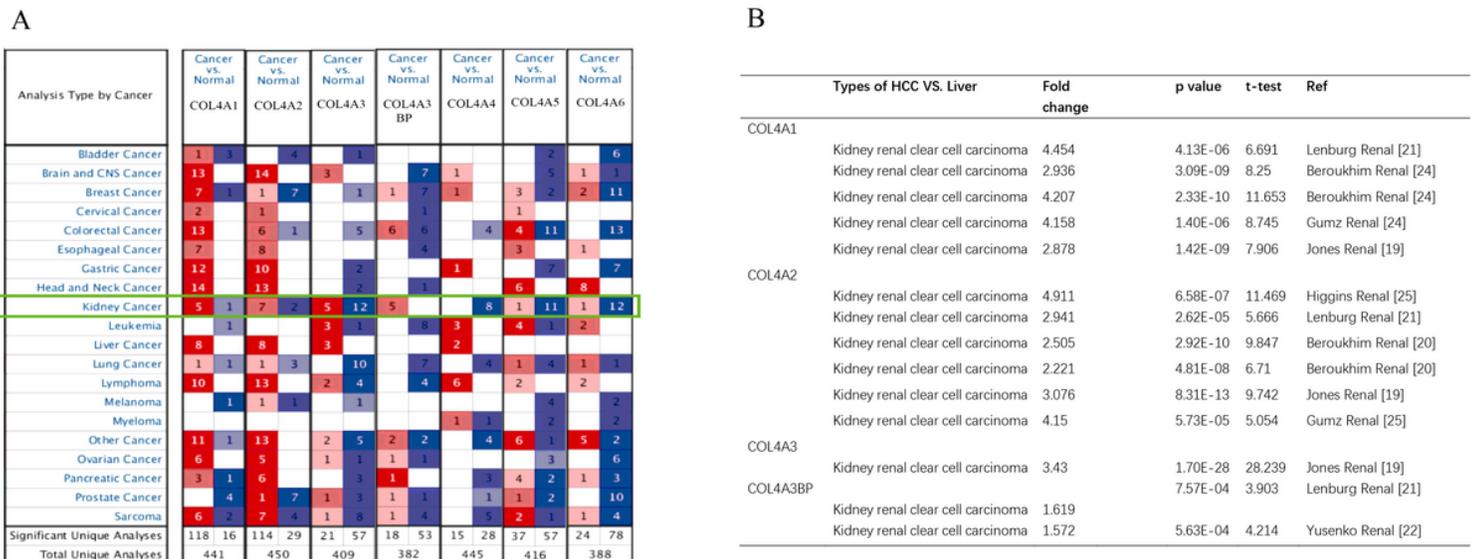


Figure 1

Transcriptional expression of COLAA family genes in kidney cancer limn (A).. Significant dungen in COLAA family gene expreadon in tramp-10m level lmiween KFIEC and normal kidney tiNsnes (8). Cu on of p •elue and fold clumgc waxas O. and1.5, l eye lOda from ONCON11.)

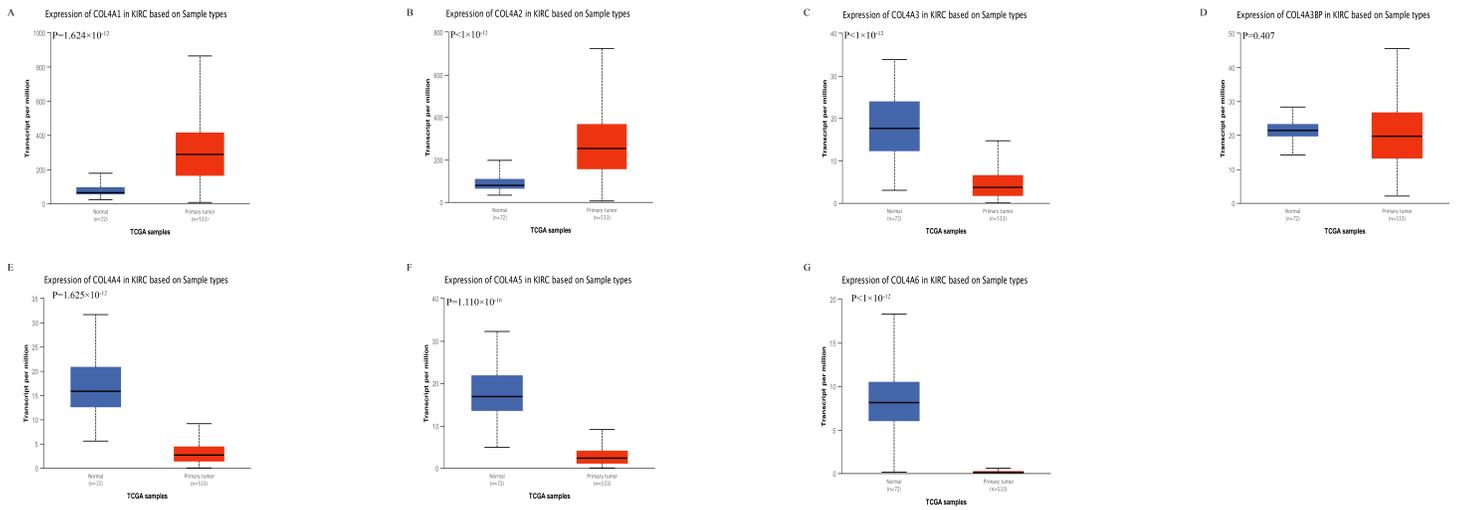


Figure 2

mRNA expression of distinct COL4A family members in Kidney renal clear cell carcinoma (KIRC) tissues and adjacent normal liver tissues (UALCAN database). mRNA expressions of five COL4As family members were found to be over-expressed in primary KIRC tissues compared to normal samples.

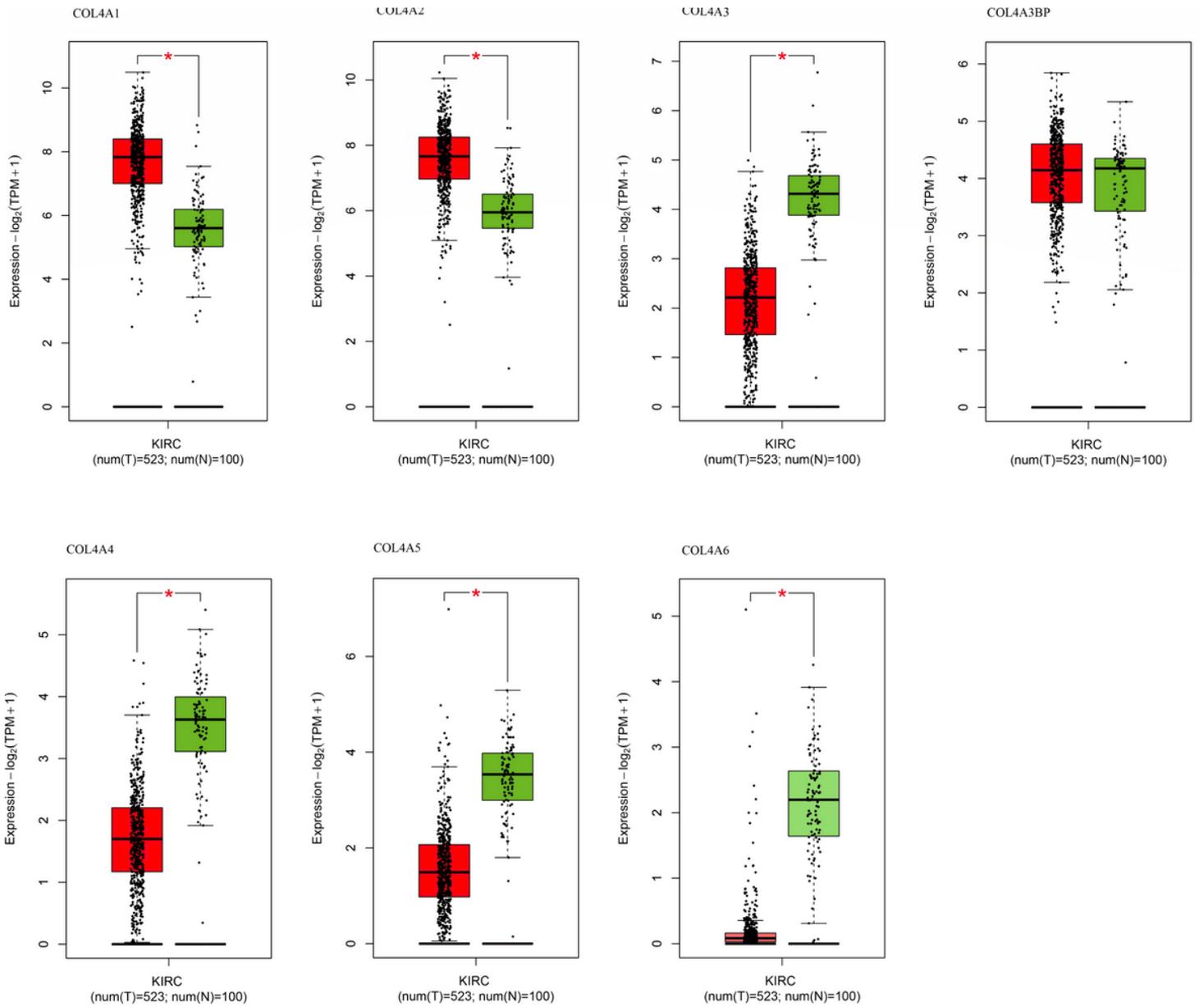


Figure 3

The tissue-wise expression of genes COL4A family in Kidney renal clear cell carcinoma (KIRC).

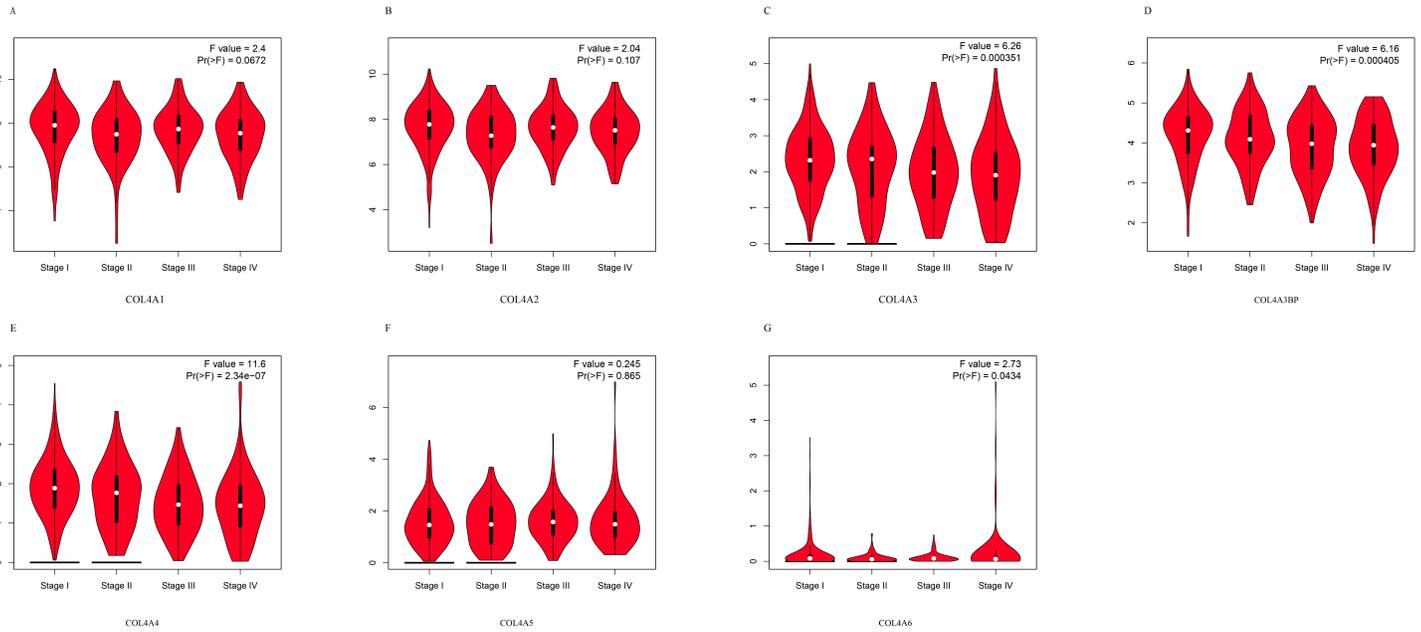


Figure 4

Correlation between COL4A family member expression and Tumor Stage in KIRC patients (Gepia). The log₂(TPM + 1) was used for log-scale.

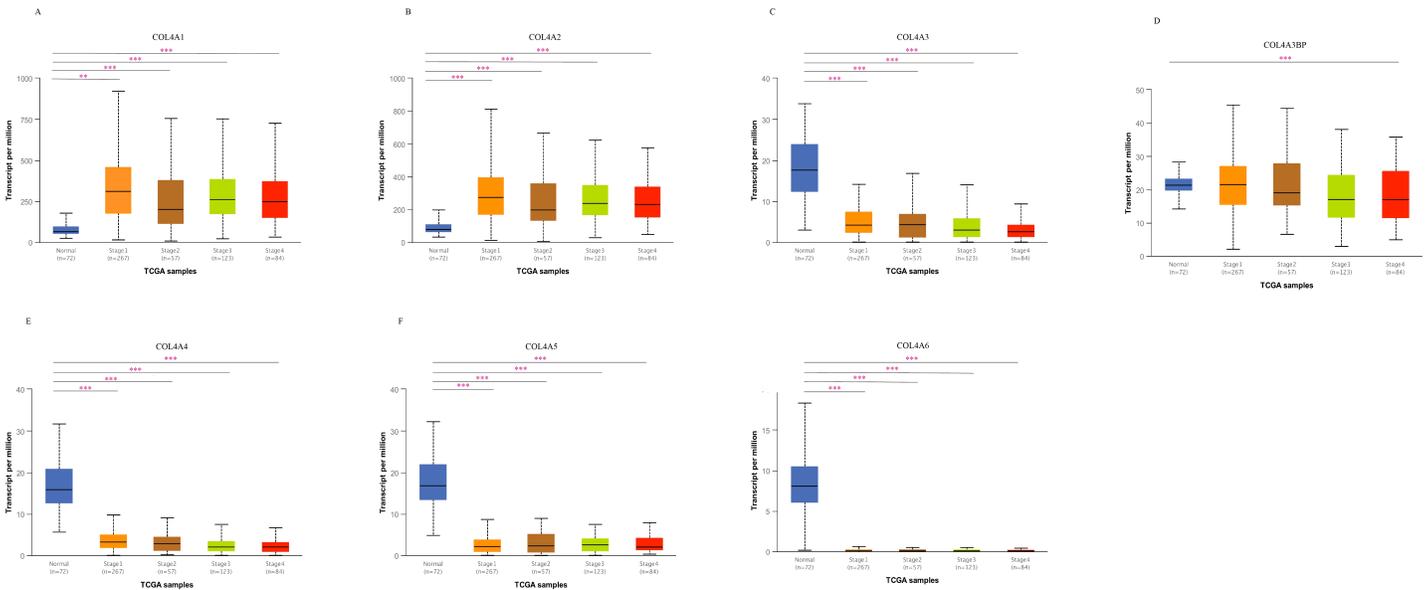


Figure 5

Relationship between mRNA expression of distinct COL4A family members and individual cancer stages of KIRC patients. mRNA expressions of 6 COL4A family members were significant correlated with patients' individual cancer stages. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

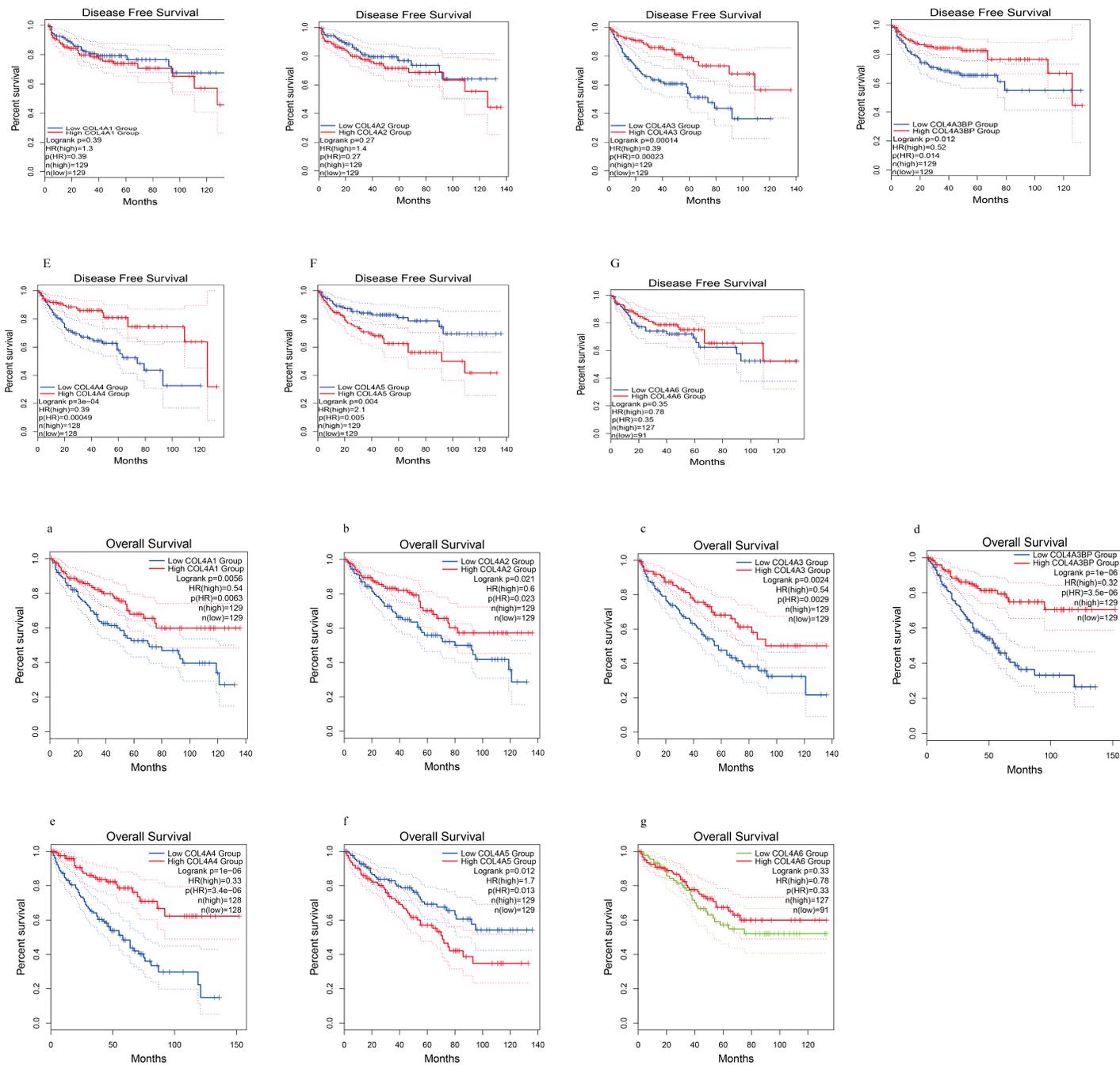


Figure 6

Prognostic value of nRNA expression of district COL4A Family members in kidney cancer patients,

Altered in 163 (32%) of queried patients/samples (510 total)

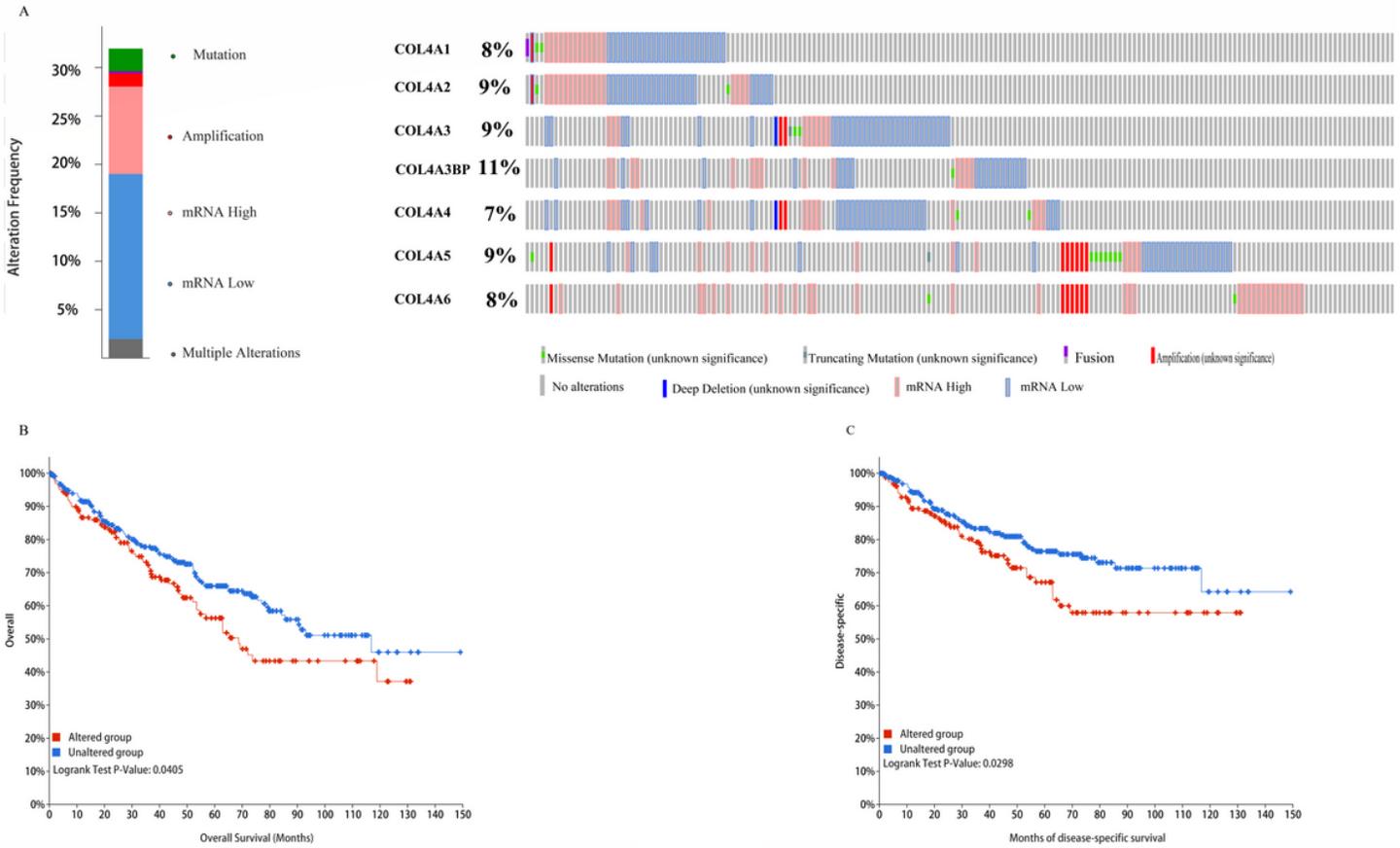
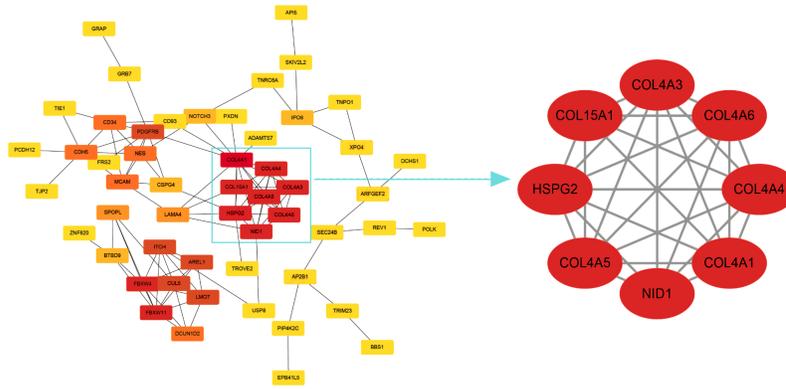


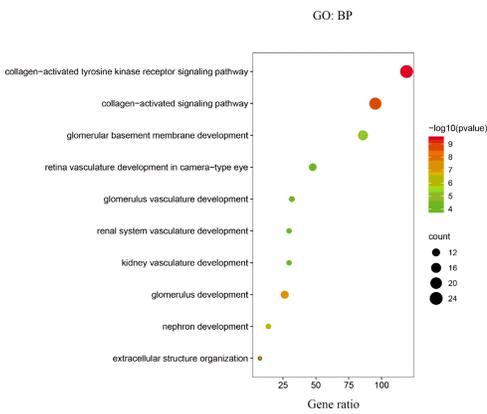
Figure 7

Genetic mutations in COL4A family genes and their association with OS and DFS of KIRC patients (cflioPortal). A high mutation rate (32%.) was observed in KIRC patients. All family members were observed genetic alterations, and their mutation rates were 8%, 9%, 9%, 11%, 7%, 9%, and 8%., respectively (A). Genetic alterations in COL4A genes were associated with shorter OS (B) and DFS (C) of HCC patients.

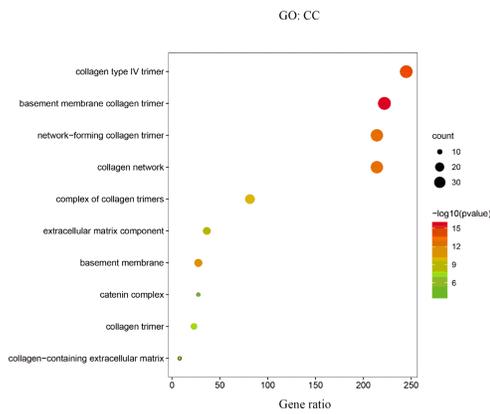
A



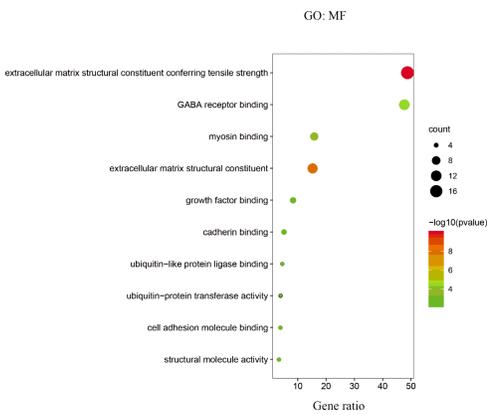
B



C



D



E

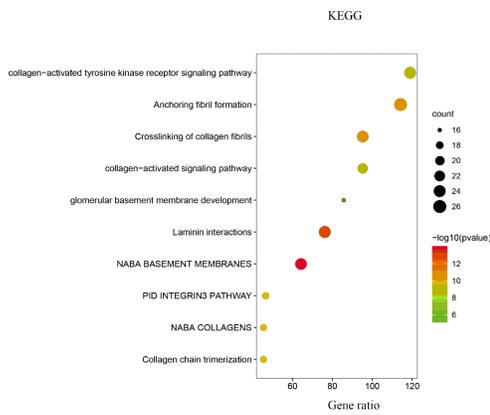


Figure 8

Predicted functions and pathways of the mutations in COL4A family genes and their 50 frequently altered neighbor genes in KIRC patients (c-BioPortal and Metascape). (A) Gene-gene interaction network for COL4A family genes and most related neighbor genes (Cytoscape). GO functional enrichment analysis predicted three main functions of COL4A family genes mutations and their 30 frequently altered neighbor genes, including biological process (B), cellular components (C) and molecular functions (D). KEGG

pathway analysis on COL4A family genes and their 30 most frequently altered neighbor genes was shown in figure E.