Key Genes Associated With Prognosis and Metastasis of Clear Cell Renal Cell Carcinoma

Tingting Zhong
Sun Yat-sen University First Affiliated Hospital https://orcid.org/0000-0002-1527-0797
Zeying Jiang
Sun Yat-sen University First Affiliated Hospital
Xiangdong Wang
Sun Yat-sen University First Affiliated Hospital
Honglei Wang
Sun Yat-sen University First Affiliated Hospital
Meiyi Song
Sun Yat-sen University First Affiliated Hospital
Wenfang Chen
Sun Yat-sen University First Affiliated Hospital
Shicong Yang (yangshc@mail.sysu.edu.cn)
Department of Pathology. The First Affiliated Hospital, Sun Yat-sen University

Research

Keywords: Clear cell renal cell carcinoma, COL6A2, COL6A3, Metastasis, P4HB, PCOLCE

Posted Date: February 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-202214/v1

Version of Record: A version of this preprint was published at PeerJ on January 4th, 2022. See the published version at https://doi.org/10.7717/peerj.12493.

Abstract

Background: Clear cell renal cell carcinoma (ccRCC) is a tumor with frequent hematogenous metastasis and is usually resistant with radiotherapy and chemotherapy. The mechanism of ccRCC metastasis still needs to be illustrated.

Material and methods: The differentially expressed genes (DEGs) of three gene expression profiles (GSE85258, GSE105288 and GSE22541) downloaded from Gene Expression Omnibus (GEO) database were analyzed by GEO2R analysis, and co-expressed DEGs among them were sorted out by the online Venn drawing tool. The co-expressed DEGs were then investigated using Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis and hub genes would be screened out based on protein-protein interaction network (PPI) established by STRING. After survival analysis performed on UALCAN website, possible key genes would be selected and then verified in ccRCC cell lines and ccRCC tissues (n=40). Statistical analysis of the above results was conducted using GraphPad Prism (Version 8.1.1).

Results: 104 co-expressed DEGs were sorted from the three profiles. KEGG pathways revealed that these genes were enriched in the extracellular matrix (ECM)-receptor interaction, focal adhesion and PI3K-Akt signaling pathway. Survival analysis showed that among 17 hub genes, patients with higher expression of Procollagen C-Endopeptidase Enhancer (PCOLCE), Prolyl 4-Hydroxylase Subunit Beta (P4HB), Collagen Type VI Alpha 2 (COL6A2), and Collagen Type VI Alpha 3 (COL6A3) had a worse survival than those with lower expression. In vitro, the mRNA expression level of PCOLCE, P4HB and COL6A2 were 3 times, and COL6A3 16 times higher in metastatic ccRCC cell line Caki-1 than in corresponding primary cell line Caki-2. Immunohistochemistry results also showed higher expression of these 4 genes in metastatic ccRCC with statistical significance.

Conclusion: PCOLCE, P4HB, COL6A2 and COL6A3 are upregulated in metastatic ccRCC and might be related to poor prognosis and distant metastases.

Highlights

1, 104 intersected co-expressed DEGs were screened out from three datasets (GSE85258, GSE105288, GSE22541) to investigate possible pathways related to prognosis and metastasis in clear cell renal cell carcinoma. GO and KEGG analyses showed DEGs mainly enriched in extracellular matrix organization, integrin binding, extracellular exosome, ECM-receptor interaction and PI3K-Akt signaling pathway.

2, Four key genes PCOLCE, P4HB, COL6A2 and COL6A3 were sorted out from 17 hub genes after survival analysis, all associated with poor survival prognosis of clear cell renal cell carcinoma.

3, In vitro, qRT-PCR assay and immunohistochemistry results showed higher expression of all 4 genes in metastatic ccRCC both in cell lines and tissues. PCOLCE and COL6A2 were found to be associated with ccRCC metastasis for the first time.

1. Introduction

CcRCC is a highly invasive malignancy that accounts for 85–90% of renal cell carcinoma[1]. About 17–30% of ccRCC patients suffered from distant metastasis by the time of diagnosis[2]. Since ccRCC is not sensitive to radiotherapy and chemotherapy, radical or partial nephrectomy is still the main treatment, with 20–40% of local recurrence or distant metastasis[3]. With the in-depth studies of ccRCC driving genes, medications towards driving gene such as hypoxia-inducible factor 2a (*HIF2a*) has been preliminarily verified and raised the hope of new targeted therapy[4]. However, the mechanism of ccRCC metastasis is still unclear. Patients with metastasis cannot get the proper treatment to improve their survival. Hence, it is extremely urgent to clarify the mechanism of ccRCC metastasis and obtain effective therapeutic targets.

Recently, microarray technology and bioinformatics analysis have facilitated the exploration of genetic alterations among various cancers. In a variety of tumors, bioinformatics analysis helps researchers to find out genes specifically related with distant metastasis. With bioinformatics analysis, researchers found sodium-potassium-chloride cotransporter 1 (NKCC1) in gastric cancer tissue was significantly higher than that in normal gastric tissue. NKCC1 could promote the proliferation, invasion, migration and epithelial to mesenchymal transition (EMT) of gastric cancer cells[5]. In oesophageal squamous cell carcinoma (ESCC), U Three Protein 14a (UTP14A) screened by bioinformatics analysis was proved to promote the proliferation and metastasis of ESCC cells by regulating PERK/eIF2a signaling pathway[6]. In order to screen out the key molecular events involved in metastatic ccRCC, we downloaded and analyzed three gene expression microarray datasets on ccRCC from GEO. By analyzing the GO and KEGG enrichment in the co-expressed DEGs, we intended to investigate pathways involving in ccRCC metastasis. We also would like to sort out key genes associated with ccRCC prognosis and metastasis through PPI network, survival analysis and in vitro verification.

2. Materials And Methods

2.1. Microarray data collection

Three gene expression datasets GSE85258[7], GSE105288[8] and GSE22541[9] were selected and downloaded from GEO (http://www.ncbi.nlm.nih.gov/geo)[10]. GSE85258 included 14 primary ccRCC samples and 15 metastatic ccRCC samples. GSE105288 included 9 primary ccRCC samples and 26 metastatic ccRCC samples. GSE22541 included 24 primary ccRCC samples and 24 metastatic ccRCC samples. GPL10558 Illumina HumanHT-12 V4.0 expression beadchip (GSE85258) and GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array (GSE105288 and GSE22541) are the microarray data of three gene datasets.

2.2. Identification of DEGs

GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r), an interactive web tool for DEGs analysis of GEO series across experimental conditions, was used to analyze the DEGs between primary ccRCC and metastatic

ccRCC within three datasets. Gene expression with |log2FoldChange| > 0.5 and $P < 0.05$ were considered as statistical significance. Next, the online Venn drawing tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) was used to obtain the intersecting co-expressed DEGs among these three datasets.

2.3. Bio-Informatic analysis of DEGs, hub genes selection and analysis

The GO datasets (biological process (BP), cellular component (CC), molecular function (MF))[11] and KEGG pathway[12] functional enrichment analysis of DEGs were conducted by DAVID (The Database for Annotation, Visualization and Integrated Discovery, http://david.ncifcrf.gov/,version6.8)[13] and $P < 0.05$ was considered to be statistically significant.

STRING (Search Tool for the Retrieval of Interacting Genes, http://string-db.org, version 11.0) was applied to predict the PPI network of DEGs[14]. An interaction with a combined score > 0.4 among DEGs was recognized as statistical significance. MCODE (Molecular Complex Detection) app of Cytoscape(version 3.7.1)[15] was used to cluster the PPI network of DEGs based on topology and to identify the closely connected area[16]. The criteria were as follows: degree cut-off = 2, node score cut-off = 0.2, Max depth = 100 and k-score = 2.

Degree ≥ 10 was set to select the hub genes. Survival analysis of hub genes was conducted on UALCAN (http://ualcan.path.uab.edu/index.html), with a $P < 0.01$ as statistically significant[17].

2.4. Cell culture

Human ccRCC cell lines Caki-1 and Caki-2 were purchased from ATCC. Cells were cultured in DMEM medium with 10% fetal bovine serum (FBS), maintained at 37°C, 5% CO_2 incubator.

2.5. qRT-PCR

TRIzol was used to extract the total RNA. Roche Transcriptor cDNA synthesis kit was employed to reversely transcript RNA to cDNA. qRT-PCR was performed on a Bio-Rad CFX-96 system (Bio-Rad), repeated in triplicate. And the sequences of primers are as follows:

Beta-Actin (F): 3’-CACCATTGGCAATGAGCGGTTC-5’

Beta-Actin (R): 3’- AGGTCTTTGCGGATGTCCACGT-5’

PCOLCE(F): 3’-GACACCTACTGCCGCTATGACT-5’

PCOLCE(R): 3’- GACGAGGAGTTCATTCCCTTCG-5’

P4HB(F): 3’-TCACCAAGGAGAACCTGCTGGA – 5’

P4HB(R): 3’-GGCAAGAACAGCAGGATGTGAG-5’

COL6A2(F): 3’-CGTGGAGACTCAGGACAGCCA-5’

COL6A2(R): 3’-CCTTTCAAGCCAAAGTCGCCTC-5’

COL6A3(F): 3’-CCTGGTGTAACTGATGCTGCCA-5’

COL6A3(R): 3’-AAGATGGCGTCCACCTTGGACT-5’

2.6. Immunohistochemistry

A series of 22 paired paraffin sections including primary ccRCC and their metastatic counterpart were collected at the First Affiliated Hospital of Sun Yat-sen University. The metastatic sites included bone (10 cases), adrenal gland (3 cases), pancreas (2 cases), liver (2 cases), lung (1 case), brain (1 case), abdominal wall (1 case), nasal cavity (1 case) and chest wall (1 case). The study was approved by the institutional ethics committee of the First Affiliated Hospital of Sun Yat-sen University.

Xylene and ethanol were used for deparaffinization and rehydration. The antigen retrieval procedure was performed with EDTA buffer in a microwave oven at 700W for 5 minutes and then 600W for 20 minutes. Slides were incubated in 3% hydrogen peroxide for 10 minutes and blocked in 5% bull serum albumin (BSA) for 10 minutes. Then slides were incubated with primary antibodies overnight at 4^{o}C (primary antibodies: *PCOLCE* (AB154261, 1:100 Abcam), *P4HB* (AB137110,1:200, Abcam), *COL6A2* (YT1035, 1:100, Immunoway), *COL6A3* (YT1036, 1:100, Immunoway)). SignalStain Boost IHC Detection Reagent (HRP, Rabbit, #8114, CST) was used as a secondary antibody for 30 minutes' incubation at room temperature. And SignalStain DAB Substrate Kit (#8059, CST) was used for staining.

IHC slides were scanned with a Digital Pathological section Scanner (KF-PRO-005) and then analyzed with Image Pro Plus to obtain the semi-quantification expression of each protein. The representative positive areas were selected in 400X magnification, which was also called the area of interesting(AOI). The integral optical density (IOD) was measured and calculated as the pixel area. The higher mean IOD of AOIs was considered to have a stronger expression.

2.7. Statistical analysis of qRT-PCR and IHC staining

Statistical analysis of qRT-PCR was analyzed by unpaired t-test, and immunohistochemistry semi-quantification was analyzed by paired t-test of GraphPad Prism (Version 8.1.1). *P*-value < 0.05 was considered as statistically significant.

3. Result

3.1. Identification of DEGs in ccRCC

A total of 996 DEGs in GSE85258, 711 DEGs in GSE105288 and 20820 DEGs in GSE2254, respectively, were obtained. Using the Venn diagram, 104 intersecting co-expressed DEGs were identified among the three datasets (Fig. 1A).

3.2. Bio-informatic analyses of DEGs and hub genes identification

GO and KEGG pathway enrichment analyses were applied in the 104 co-expressed DEGs with DAVID online. The GO results revealed that in biological process (BP), DEGs were mainly enriched in extracellular matrix organization, cell adhesion, collagen catabolic process, proteolysis, oxidation-reduction process, positive regulation of cell proliferation, etc. (Fig. 1B). In molecular function (MF), DEGs were predominantly enriched in integrin binding, SMAD binding, collagen binding, cell adhesion molecule binding, heparin binding, protease binding, extracellular matrix structural constituent, peptidase activator activity, and serine-type peptidase activity (Fig. 1C). As for cellular component (CC), DEGs were enriched in extracellular exosome, extracellular space, extracellular region, extracellular matrix, proteinaceous extracellular matrix, collagen trimer, endoplasmic reticulum lumen and basement membrane (Fig. 1D). KEGG pathway analysis indicated that the DEGs were enriched in focal adhesion, ECM-receptor interaction, protein digestion and absorption, PI3K-Akt signaling pathway, platelet activation, complement and coagulation cascades, etc. (Fig. 1E).

3.3. PPI network construction,hub gene selection and key genes identification

STRING online database demonstrated the PPI of the DEGs. Among them, the most significant module containing 17 hub genes was established by the MCODE app of Cytoscape (Fig. 2A, 2B). GO and KEGG pathway enrichment analyses of 17 hub genes were showed in Table 1. From GO enrichment results, 17 hub genes were mainly enriched in collagen fibril organization, cellular response to amino acid stimulus and protein heterotrimerization in BP, proteinaceous extracellular matrix, collagen trimer and extracellular space in CC, and extracellular matrix structural constituent, platelet-derived growth factor binding and peptidase activator activity in MF. KEGG pathways were mainly enriched Protein digestion and absorption, ECM-receptor interaction, focal adhesion and PI3K-Akt signaling pathway (Table 1).

Table 1
GO and KEGG pathway enrichment analysis of 17 hub DEGs

Term	Description	Count in gene set	Gene ratio (%)	*P*-Value
GOTERM_BP_DIRECT				
GO:0030199	Collagen fibril organization	6	21%	3.59E-11
GO:0071230	Cellular response to amino acid stimulus	5	17%	6.25E-08
GO:0070208	Protein heterotrimerization	3	10%	2.93E-05
GO:0043589	Skin morphogenesis	2	7%	0.006565634
GO:0043588	Skin development	2	7%	0.023266414
GO:0001568	Blood vessel development	2	7%	0.026024548
GO:0001501	Skeletal system development	2	7%	0.047832184
GO:0007179	Transforming growth factor beta receptor signaling pathway	2	7%	0.047832184
GOTERM_CC_DIRECT				
GO:0005578	Proteinaceous extracellular matrix	9	31%	7.81E-13
GO:0005581	Collagen trimer	7	24%	5.62E-12
GO:0005615	Extracellular space	10	35%	1.82E-08
GO:0070062	Extracellular exosome	9	31%	1.97E-04
GO:0005584	Collagen type I trimer	2	7%	0.001810247
GO:0042383	Sarcolemma	2	7%	0.049551877
GOTERM_MF_DIRECT				
GO:0005201	Extracellular matrix structural constituent	5	17%	1.75E-08
GO:0048407	Platelet-derived growth factor binding	2	7%	0.002518122
GO:0016504	Peptidase activator activity	2	7%	0.006702184
GO:0005518	Collagen binding	2	7%	0.012533069
KEGG_PATHWAY				
bta04974	Protein digestion and absorption	8	28%	1.91E-12
bta04512	ECM-receptor interaction	7	24%	3.98E-10
bta04510	Focal adhesion	7	24%	7.80E-08
bta04151	PI3K-Akt signaling pathway	7	24%	1.62E-06

Term	Description	Count in gene set	Gene ratio (%)	*P*-Value
bta05146	Amoebiasis	5	17%	9.00E-06
bta04611	Platelet activation	5	17%	1.48E-05
GO, Gene Ontology; MF, Molecular Function; CC, cellular component; BP, biological process; KEGG, Kyoto Encyclopedia of Genes and Genomes.				

4 key genes, *PCOLCE*, *P4HB*, *COL6A2* and *COL6A3* were sorted out with *P*-value < 0.01 among the 17 hub genes after survival analysis performed on the UALCAN website. Data showed that patients with higher expression of *PCOLCE*, *P4HB*, *COL6A2* and *COL6A3* had worse survival than those with lower expression (Fig. 2C-2F).

3.4. In vitro testification of key genes

qRT-PCR showed *PCOLCE*, *P4HB*, *COL6A2* and *COL6A3* were all upregulated in Caki-1 cell lines when compared with its primary counterpart Caki-2 cell line ($P < 0.05$). (Fig. 3A). Protein expression of these four genes was examined with immunohistochemistry using 22 pairs of primary ccRCC and their corresponding metastatic counterpart tissue. Mean IOD results of semi-quantification showed that in most paired tissues, higher expressions of all 4 proteins were observed in metastatic ccRCC tissue rather than primary sites($P < 0.05$). In all, higher expressions in metastatic counterparts were observed in 19 out of 22 pairs regarding *COL6A2* and *PCOLCE*, 18 out of 22 pairs regarding *COL6A3* and *P4HB* expression. (Fig. 3B).

4. Discussion

CcRCC was believed to raise from the proximal tubular epithelium with aggressive behavior that was unparallel with its mild appearance. After surgery, the median survival time of patients with metastasis was only 13 months, and the 5-year survival rate was merely 12.3%[18]. As one of the main genetic events of ccRCC pathogenesis, the mutant *VHL* gene pathway has become the target of treatment towards ccRCC[4]. Hopefully, with the help of bio-informatic data management and further verification, potential key genes could be identified and act as new targets for ccRCC patients with metastasis.

In our study, bioinformatic analysis from the 104 co-expressed DEGs showed concrete pathways that might be related to ccRCC metastasis. GO analysis showed DEGs mainly distributed in extracellular space or exosome and were enriched in pathways with extracellular matrix organization, cell adhesion, integrin binding and extracellular matrix structural constituent, etc. KEGG pathway analysis revealed these DEGs enriched in focal adhesion, ECM-receptor interaction and PI3K-Akt signaling pathway. Exosomes, by carrying proteins, ceramide, mRNA or micro-RNA, can induce metastasis through participation in EMT, mesenchymal-epithelial reverting transition (MErT), and neoangiogenesis[19]. For example, in ccRCC cell lines, exosomes containing miR-19b-3p can initiate EMT and accelerate tumor

metastasis by reducing PTEN expression[20]. EMT is essential in tumor metastasis through processes such as disorienting cell polarity, disconnection with the basement membrane, acquiring plasticity, invasiveness and anti-apoptosis[21]. Studies have shown that the PI3K-Akt signaling pathway might induce the EMT process directly or through cross-talk with the Wnt/b-catenin signaling pathway, TGF-b, or NF-kB and initiate distant metastasis[22]. In ccRCC, the phosphorylation of AKT, which was a key kinase of the PI3K signaling pathway was reported to be related to EMT[23]. Our results also showed DEGs enriched in the PI3K signaling pathway, suggesting that PI3K signaling pathway-mediated EMT was likely to participate in the metastasis of ccRCC. To sum up, our analysis results have indicated various processes that might be associated with ccRCC metastases.

During the effort of establishing key genes, we first obtained 17 hub genes from the PPI network analysis. Using survival analysis, *PCOLCE*, *P4HB*, *COL6A2* and *COL6A3* were found related to ccRCC patient survival, which were then further tested with qRT-PCR and IHC staining. We found that these 4 genes presented higher expression in the metastatic counterpart of ccRCC both in cell lines and in the paired ccRCC tissues. Among them, *P4HB* and *COL6A3* have been reported to be associated with ccRCC, whereas *PCOLCE* and *COL6A2* were found to be associated with ccRCC for the first time. By making primary exploration, we thought these 4 genes might play an important role during the metastases of ccRCC.

To our knowledge, this was the first time that *PCOLCE* was found to be related to ccRCC. *PCOLCE* is a secreted glycoprotein that participates in the maturation of procollagen and ECM reconstruction by enhancing the activity of bone morphogenetic protein-1 (*BMP-1*)[24–26]. *PCOLCE* was overexpressed in osteosarcoma, which could promote lung metastasis via twist family bHLH transcription factor 1 (*TWIST1*). Reducing *PCOLCE* expression could prevent the migration, invasion and lung metastasis of osteosarcoma cells[27]. *BMP-1* was reported to be up-regulated in gastric cancer associated with poor survival and distant metastasis[28]. Importantly, *BMP-1* is elevated in ccRCC and promoted proliferation, migration, and invasion in ccRCC cell lines[29]. Our study showed *PCOLC*E, as an active enhancer of *BMP-1*, was higher in metastatic ccRCC, which raised the hypothesis that *PCOLCE* might be involved in ccRCC metastasis through regulation on *BMP-1.* Therefore, *PCOLCE* might act as a potential research target on ccRCC metastasis and more data was needed for further research.

P4HB, a beta subunit of Prolyl 4-hydroxylase, might help cancer cells survive from apoptosis induced by endoplasmic reticulum (ER) stress[30, 31]. *P4HB* promoted proliferation, invasion, migration and angiogenesis in glioma through mitogen-activated protein kinase (MAPK) signaling pathway[32]. In gastric cancer, *P4HB* was proved to play an important role on regulating invasion and migration in the hypoxia-inducible factor 1a (*HIF1a*) network[33]. Previous studies showed that overexpression of *P4HB* in ccRCC was associated with poor outcome[34]. In our study, we further found that *P4HB* was specifically elevated in metastatic ccRCC. Hence, we considered *P4HB* might play a key role in modulating ccRCC metastasis.

Also, *COL6A2* was found to be related to ccRCC progression for the first time. *COL6A1-3* were three subunits of collagen VI[35, 36]. Strong evidence had demonstrated *COL6A1-3*'s role in promoting the progression of tumors[37]. *COL6A2* was dysregulated and associated with poor prognosis by involving in TGFβ1 pathway in serous ovarian cancer[38, 39]. *COL6A3* may take part in the tumorigenesis and progression of cholangiocarcinoma through E2F1/LMCD1-AS1/miR-345-5p/COL6A3 axis and might act as a prognostic factor for pancreatic cancer[40]. Up-regulation of *COL6A1* in ccRCC was associated with poor prognosis of patients and promoted tumor growth in vivo[41]. *COL6A3* was reported to be overexpressed in metastatic renal cell carcinoma (RCC) and associated with poor survival[42]. To date, there is no report mentioning the relation between *COL6A2* and ccRCC. As *COL6A2* and *COL6A3* were identified as hub genes in our results, an in-depth research is worth trying to clarify their mechanism on ccRCC metastasis.

With meaningful findings in our study as to establishing 4 possible key genes related to ccRCC metastasis, there are many limitations. Firstly, we only conducted 22 paired ccRCC samples and 2 cell lines to verify the bio-informatic findings, the size of which needs to be enlarged to obtain more reliable results. Secondly, the mechanism of four key genes in regulating ccRCC cells was still unclear. Further exploration should be conducted to clarify the function and potential mechanism of the four key genes both in vitro and in vivo.

5. Conclusion

In conclusion, our studies established 4 key genes (*PCOLCE*, *P4HB*, *COL6A2* and *COL6A3*) that were related to patient survival and might be involved in ccRCC metastasis. These 4 key genes may be involved in the ECM-receptor interaction, focal adhesion and PI3K-Akt signaling pathway, etc. In vitro, experiments have shown increased expression of these 4 genes in metastatic ccRCC both in cell lines and tissues. Further investigation was worth taking out to evaluate the significance and mechanism of these key genes.

6. List Of Abbreviations

AOI, area of interesting; BMP-1, bone morphogenetic protein-1; BP, biological process; BSA, bull serum albumin; ccRCC, clear cell renal cell carcinoma; CC,cellular component; COL6A2, collagen Type VI Alpha 2; COL6A3, collagen Type VI Alpha 3; DEGs, differentially expressed genes; ER, endoplasmic reticulum; ECM, extracellular matrix; EMT, epithelial to mesenchymal transition; ESCC, oesophageal squamous cell carcinoma; FBS, fetal bovine serum; GEO, gene Expression Omnibus; GO, gene Ontology; HIF2a, hypoxia-inducible factor 2a; HIF1a, hypoxia-inducible factor 1a; IHC, immunohistochemistry; IOD, integral optical density; KEGG, Kyoto Encyclopedia of Genes and Genomes; MF, molecular function; MCODE, molecular Complex Detection; MErT, mesenchymal-epithelial reverting transition; MAPK, mitogen-activated protein kinase; NKCC1, sodium-potassium-chloride cotransporter 1; PPI, protein-protein interaction network; PCOLCE, procollagen C-Endopeptidase Enhancer; P4HB, prolyl 4-Hydroxylase Subunit Beta; RCC, renal cell carcinoma; TWIST, twist family bHLH transcription factor 1; UTP14A, U Three Protein 14a

7. Declarations

7.1. Ethics approval and consent to participate:

The study was approved by the institutional ethics committee of the First Affiliated Hospital of Sun Yat-sen University.

7.2. Data availability statement:

The datasets that support the findings of this study are openly available in GEO at https://doi.org/10.1093/annonc/mdw652, https://doi.org/10.1158/1541-7786.Mcr-17-0636, https://doi.org/10.1002/ijc.27419. In vitro experimental data that support the findings of this study are available from the corresponding author upon reasonable request.

7.3. Competing interests:

All the authors declared that we do not have any commercial or associative interest conflicts with this work.

7.4. Funding:

Our study is funded by National Natural Science Foundation of China (No. 8177100503).

7.5. Authors' contributions:

Designed and conceived the experiments by Shicong Yang and Tingting Zhong; Data analysis: Tingting Zhong and Wenfang Chen; Performed The experiments: Tingting Zhong, Zeying Jiang, Xiangdong Wang, Honglei Wang; Tissue paraffin section preparation: Meiyi Song; The writing original draft preparation be completed: Tingting Zhong and Shicong Yang. All authors consented this publication.

7.6 Acknowledgements:

We thank the Nephropathy Laboratory of the First Affiliated Hospital of Sun Yat-sen University for providing instruments for our experiments in vitro.

References

1. J. Chen,Y. Shu,Q. Yu, et al., MicroRNA-645 promotes cell metastasis and proliferation of renal clear cell carcinoma by targeting GK5, Eur Rev Med Pharmacol Sci (2017) 4557-4565.
2. R.J. Motzer,N.H. BanderD.M. Nanus, Renal-cell carcinoma, N Engl J Med (1996) 865-75.
3. J. Teng,Y. Gao,M. Chen, et al., Prognostic value of clinical and pathological factors for surgically treated localized clear cell renal cell carcinoma, Chin Med J (Engl) (2014) 1640-4.
4. W. Chen,H. Hill,A. Christie, et al., Targeting renal cell carcinoma with a HIF-2 antagonist, Nature (2016) 112-117.

5. J.F. Wang,K. Zhao,Y.Y. Chen, et al., NKCC1 promotes proliferation, invasion and migration in human gastric cancer cells via activation of the MAPK-JNK/EMT signaling pathway, J Cancer (2021) 253-263.
6. K.K. Li,C.Y. Mao,Q. Ma, et al., U three protein 14a (UTP14A) promotes tumour proliferation and metastasis via the PERK/eIF2a/GRP78 signalling pathway in oesophageal squamous cell carcinoma, J Cancer (2021) 134-140.
7. T.H. Ho,D.J. Serie,M. Parasramka, et al., Differential gene expression profiling of matched primary renal cell carcinoma and metastases reveals upregulation of extracellular matrix genes, Ann Oncol (2017) 604-610.
8. H.Y. Nam,D.S. Chandrashekar,A. Kundu, et al., Integrative Epigenetic and Gene Expression Analysis of Renal Tumor Progression to Metastasis, Mol Cancer Res (2019) 84-96.
9. D. Wuttig,S. Zastrow,S. Füssel, et al., CD31, EDNRB and TSPAN7 are promising prognostic markers in clear-cell renal cell carcinoma revealed by genome-wide expression analyses of primary tumors and metastases, Int J Cancer (2012) E693-704.
10. R. Edgar,M. DomrachevA.E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository, Nucleic Acids Res (2002) 207-10.
11. M. Ashburner,C.A. Ball,J.A. Blake, et al., Gene ontology: tool for the unification of biology. The Gene Ontology Consortium, Nat Genet (2000) 25-9.
12. M. Kanehisa,S. Goto,M. Furumichi, et al., KEGG for representation and analysis of molecular networks involving diseases and drugs, Nucleic Acids Res (2010) D355-60.
13. G. Dennis, Jr.,B.T. Sherman,D.A. Hosack, et al., DAVID: Database for Annotation, Visualization, and Integrated Discovery, Genome Biol (2003) P3.
14. D. Szklarczyk,A. Franceschini,S. Wyder, et al., STRING v10: protein-protein interaction networks, integrated over the tree of life, Nucleic Acids Res (2015) D447-52.
15. M.E. Smoot,K. Ono,J. Ruscheinski, et al., Cytoscape 2.8: new features for data integration and network visualization, Bioinformatics (2011) 431-2.
16. W.P. Bandettini,P. Kellman,C. Mancini, et al., MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study, J Cardiovasc Magn Reson (2012) 83.
17. D.S. Chandrashekar,B. Bashel,S.A.H. Balasubramanya, et al., UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses, Neoplasia (2017) 649-658.
18. K.N. Babaian,M.M. Merrill,S. Matin, et al., Partial nephrectomy in the setting of metastatic renal cell carcinoma, J Urol (2014) 36-42.
19. T.B. Steinbichler,J. Dudás,H. Riechelmann, et al., The role of exosomes in cancer metastasis, Semin Cancer Biol (2017) 170-181.
20. L. Wang,G. Yang,D. Zhao, et al., CD103-positive CSC exosome promotes EMT of clear cell renal cell carcinoma: role of remote MiR-19b-3p, Mol Cancer (2019) 86.

21. L.J. Talbot,S.D. BhattacharyaP.C. Kuo, Epithelial-mesenchymal transition, the tumor microenvironment, and metastatic behavior of epithelial malignancies, Int J Biochem Mol Biol (2012) 117-36.
22. W. Xu,Z. YangN. Lu, A new role for the PI3K/Akt signaling pathway in the epithelial-mesenchymal transition, Cell Adh Migr (2015) 317-24.
23. S.T. Wu,G.H. Sun,C.Y. Hsu, et al., Tumor necrosis factor-α induces epithelial-mesenchymal transition of renal cell carcinoma cells via a nuclear factor kappa B-independent mechanism, Exp Biol Med (Maywood) (2011) 1022-9.
24. E. Kessler,A.P. MouldD.J. Hulmes, Procollagen type I C-proteinase enhancer is a naturally occurring connective tissue glycoprotein, Biochem Biophys Res Commun (1990) 81-6.
25. C. Moali,B. Font,F. Ruggiero, et al., Substrate-specific modulation of a multisubstrate proteinase. C-terminal processing of fibrillar procollagens is the only BMP-1-dependent activity to be enhanced by PCPE-1, J Biol Chem (2005) 24188-94.
26. Q. Zhu,W. Guo,S. Zhang, et al., Synergistic effect of PCPE1 and sFRP2 on the processing of procollagens via BMP1, FEBS Lett (2019) 119-127.
27. S. Wang,L. Zhong,Y. Li, et al., Up-regulation of PCOLCE by TWIST1 promotes metastasis in Osteosarcoma, Theranostics (2019) 4342-4353.
28. Y.Y. Hsieh,S.Y. Tung,H.Y. Pan, et al., Upregulation of bone morphogenetic protein 1 is associated with poor prognosis of late-stage gastric Cancer patients, BMC Cancer (2018) 508.
29. W. Xiao,X. Wang,T. Wang, et al., Overexpression of BMP1 reflects poor prognosis in clear cell renal cell carcinoma, Cancer Gene Ther (2020) 330-340.
30. L. Pajunen,T.A. Jones,A. Goddard, et al., Regional assignment of the human gene coding for a multifunctional polypeptide (P4HB) acting as the beta-subunit of prolyl 4-hydroxylase and the enzyme protein disulfide isomerase to 17q25, Cytogenet Cell Genet (1991) 165-8.
31. P.E. Lovat,M. Corazzari,J.L. Armstrong, et al., Increasing melanoma cell death using inhibitors of protein disulfide isomerases to abrogate survival responses to endoplasmic reticulum stress, Cancer Res (2008) 5363-9.
32. S. Sun,K.M.Y. Kiang,A.S.W. Ho, et al., Endoplasmic reticulum chaperone prolyl 4-hydroxylase, beta polypeptide (P4HB) promotes malignant phenotypes in glioma via MAPK signaling, Oncotarget (2017) 71911-71923.
33. J. Zhang,S. Guo,Y. Wu, et al., P4HB, a Novel Hypoxia Target Gene Related to Gastric Cancer Invasion and Metastasis, Biomed Res Int (2019) 9749751.
34. Z. Zhu,A. He,T. Lv, et al., Overexpression of P4HB is correlated with poor prognosis in human clear cell renal cell carcinoma, Cancer Biomark (2019) 431-439.
35. H. HessleE. Engvall, Type VI collagen. Studies on its localization, structure, and biosynthetic form with monoclonal antibodies, J Biol Chem (1984) 3955-61.

36. D. Weil,M.G. Mattei,E. Passage, et al., Cloning and chromosomal localization of human genes encoding the three chains of type VI collagen, Am J Hum Genet (1988) 435-45.
37. P. Chen,M. CesconP. Bonaldo, Collagen VI in cancer and its biological mechanisms, Trends Mol Med (2013) 410-7.
38. D.J. Cheon,Y. Tong,M.S. Sim, et al., A collagen-remodeling gene signature regulated by TGF-β signaling is associated with metastasis and poor survival in serous ovarian cancer, Clin Cancer Res (2014) 711-23.
39. C. Svoronos,G. Tsoulfas,M. Souvatzi, et al., Prognostic value of COL6A3 in pancreatic adenocarcinoma, Ann Hepatobiliary Pancreat Surg (2020) 52-56.
40. J. Yu,B. Zhang,H. Zhang, et al., E2F1-induced upregulation of long non-coding RNA LMCD1-AS1 facilitates cholangiocarcinoma cell progression by regulating miR-345-5p/COL6A3 pathway, Biochem Biophys Res Commun (2019) 150-155.
41. F. Wan,H. Wang,Y. Shen, et al., Upregulation of COL6A1 is predictive of poor prognosis in clear cell renal cell carcinoma patients, Oncotarget (2015) 27378-87.
42. C. Li,T. Shao,G. Bao, et al., Identification of potential core genes in metastatic renal cell carcinoma using bioinformatics analysis, Am J Transl Res (2019) 6812-6825.

Figures

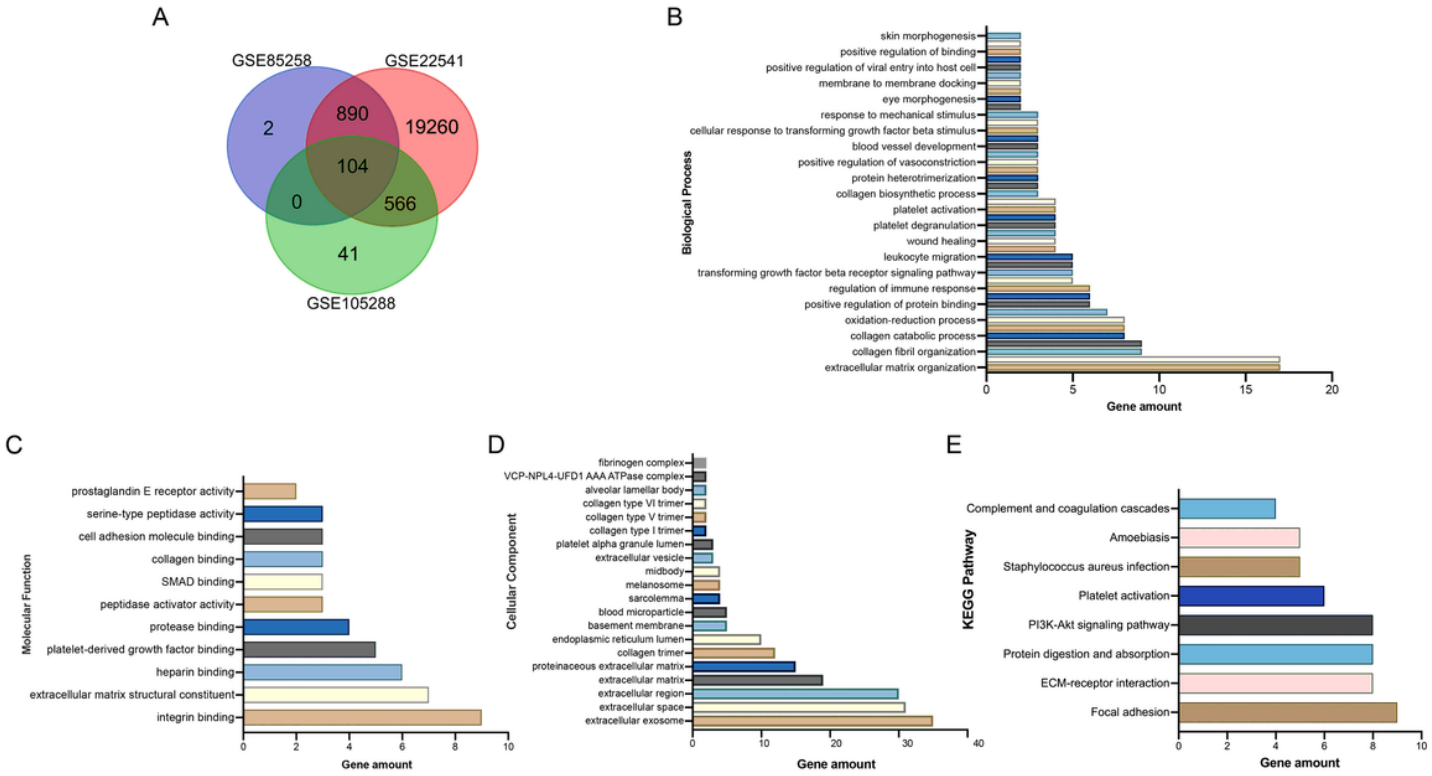

Figure 1

Venn diagram, GO enrichment analysis and KEGG pathway analysis (P-value <0.05). (A) DEGs were selected with P-value < 0.05 and |log2 FoldChange | > 0.5 among the datasets GSE85258, GSE105288 and GSE22541. Venn diagram showed an overlap of 104 genes. (B) All BP results of 104 DEGs. (C) All CC results of 104 DEGs. (D) All MF results of 104 DEGs. (E) All KEGG pathway enrichment results of the 104 DEGs. Graphs of the GO enrichment and KEGG pathway enrichment results were generated by GraphPad Prism (Version 8.1.1).

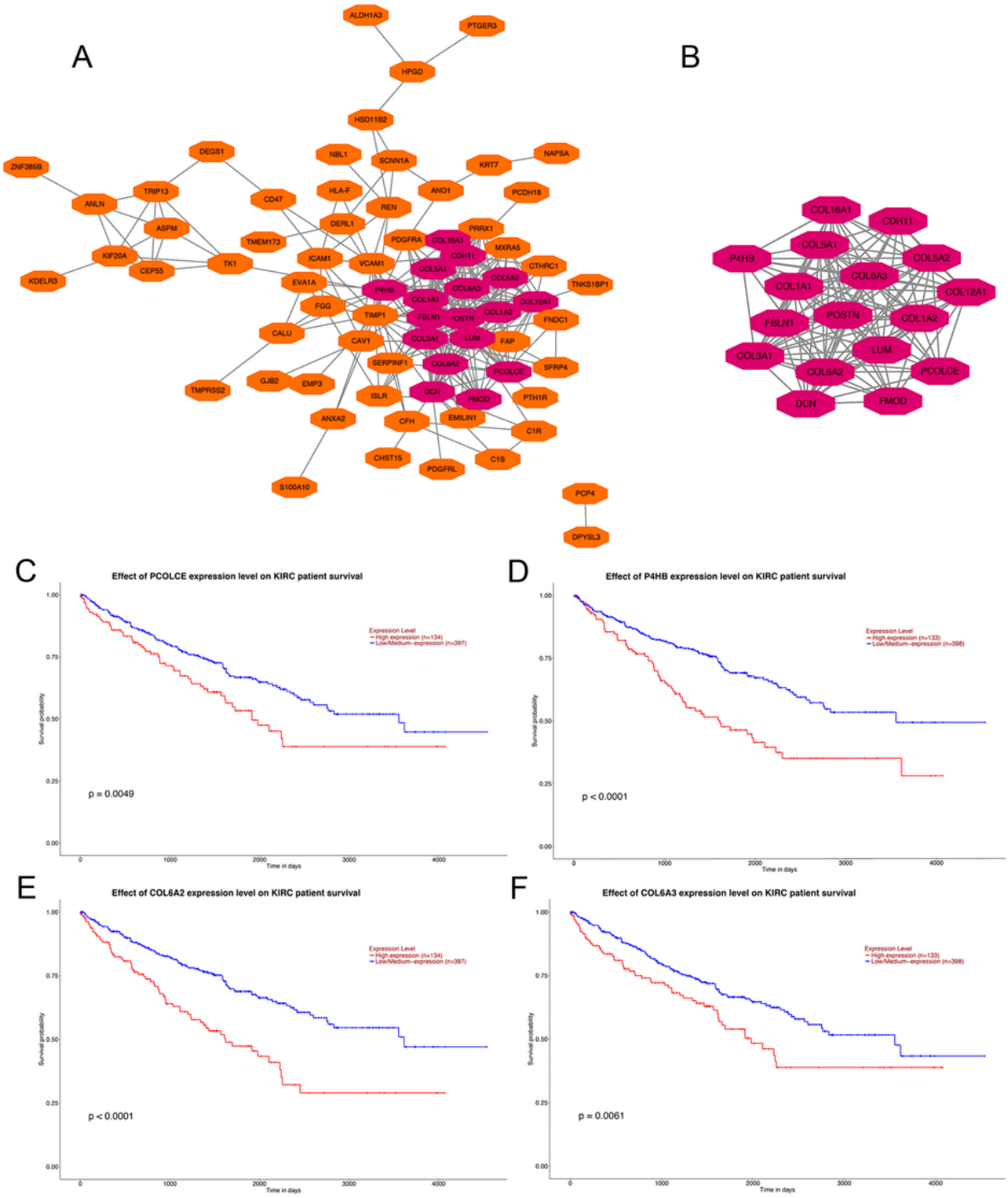

Figure 2

PPI network of DEGs, hub genes selection, and survival analysis of 4 key genes. (A) PPI network of 104 DEGs, the representation of nodes and edges are protein and the interaction of the proteins. (B) 17 hub genes of DEGs were obtained with module analysis using the MCODE app. (P-value <0.01, KIRC: kidney renal clear cell carcinoma). (C) Survival analysis of PCOLCE. (D) Survival analysis of P4HB. (E) Survival analysis of COL6A2. (F) Survival analysis of COL6A3.

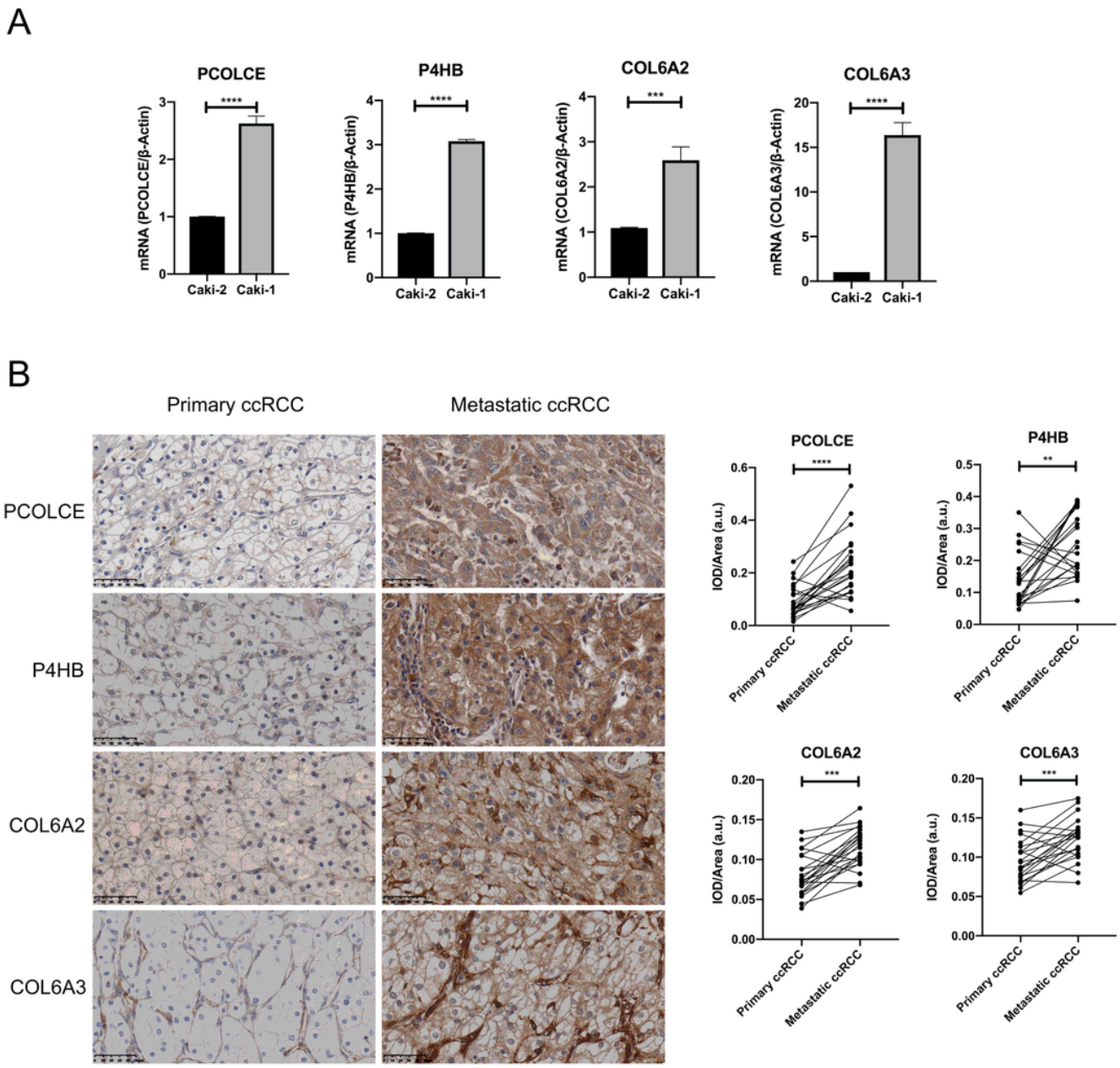

Figure 3

Expression of key genes in ccRCC and corresponding metastatic ccRCC. (A) the qRT-PCR analysis showed the expression of PCOLCE, P4HB, COL6A2 and COL6A3 in primary ccRCC cell line Caki-2 and its corresponding metastatic ccRCC cell line Caki-1. (***P < 0.001, ****P < 0.0001) (B) Representative IHC staining images of PCOLCE, P4HB, COL6A2 and COL6A3 in human primary ccRCC and corresponding metastatic ccRCC (22 paired cases) on the left. And the semi-quantitative results on the right. 400x. (**P < 0.01, *** P < 0.001, ****P < 0.0001).