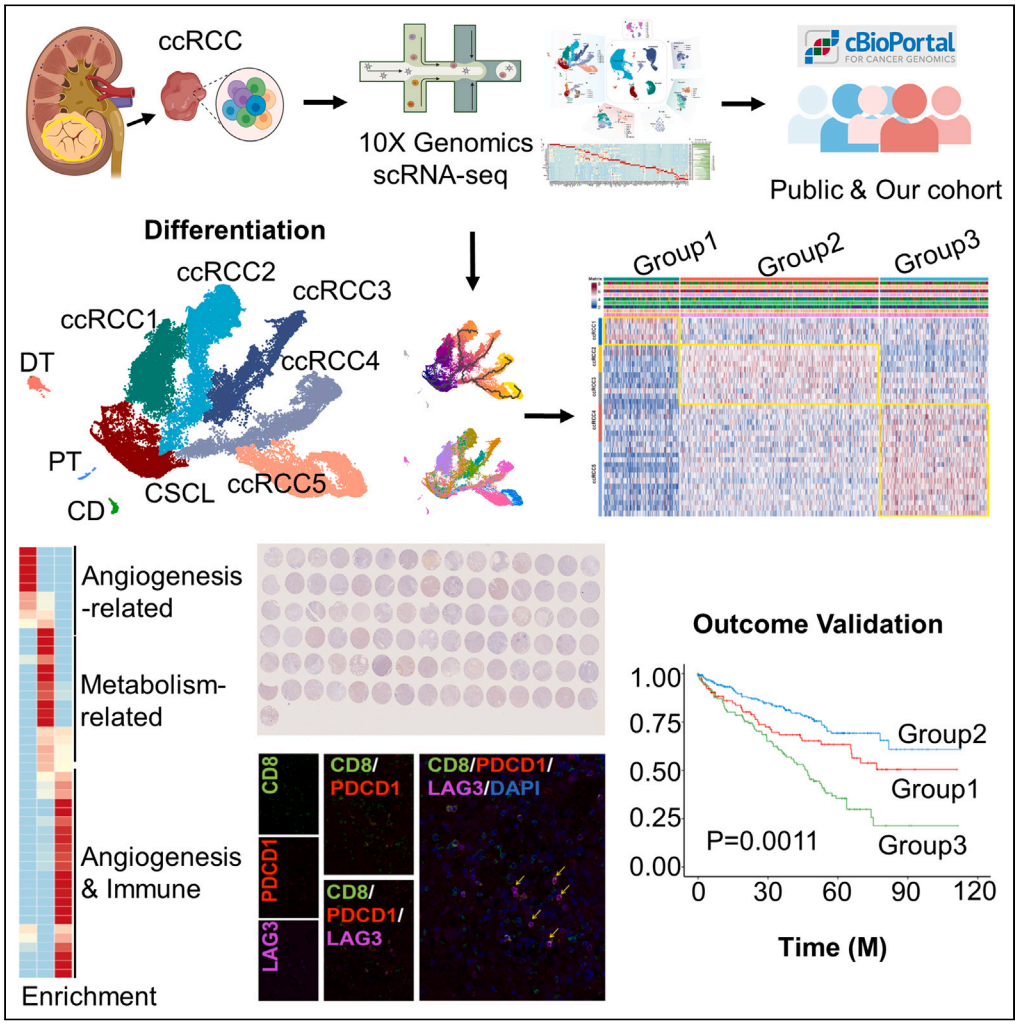
iScience

CellPress
OPEN ACCESS

Article

Molecular subtyping and characterization of clear cell renal cell carcinoma by tumor differentiation trajectories



Xiu-wu Pan, Wen-jin Chen, Da Xu, ..., Si-shun Gan, Wang Zhou, Xin-gang Cui

gansishun20101111@163.com (S.-s.G.)
brilliant212@163.com (W.Z.)
cuixingang@xinhuamed.com.cn (X.-g.C.)

Highlights

Differentiation trajectories from scRNA-seq were used for molecular subtyping

Three ccRCC subgroups are representative for outcomes and therapies

Transcriptional landscape of cancer stem cells in ccRCC

Two terminally exhausted states of CD8+T cells show different roles in treatment

Pan et al., iScience 26, 108370
December 15, 2023 © 2023 The Authors.
https://doi.org/10.1016/j.isci.2023.108370

iScience

CellPress
OPEN ACCESS

Article

Molecular subtyping and characterization of clear cell renal cell carcinoma by tumor differentiation trajectories

Xiu-wu Pan,[1,4] Wen-jin Chen,[1,2,4] Da Xu,[2,4] Wen-bin Guan,[3,4] Lin Li,[2] Jia-xin Chen,[1] Wei-jie Chen,[2] Ke-qin Dong,[1] Jian-qing Ye,[1] Si-shun Gan,[2,*] Wang Zhou,[1,*] and Xin-gang Cui[1,5,*]

SUMMARY

Previous bulk RNA sequencing or whole genome sequencing on clear cell renal cell carcinoma (ccRCC) subtyping mainly focused on ccRCC cell origin or the complex tumor microenvironment (TME). Based on the single-cell RNA sequencing (scRNA-seq) data of 11 primary ccRCC specimens, cancer stem-cell-like subsets could be differentiated into five trajectories, whereby we further classified ccRCC cells into three groups with diverse molecular features. These three ccRCC subgroups showed significantly different outcomes and potential targets to tyrosine kinase inhibitors (TKIs) or immune checkpoint inhibitors (ICIs). Tumor cells in three differentiation directions exhibited distinct interactions with other subsets in the ccRCC niches. The subtyping model was examined through immunohistochemistry staining in our ccRCC cohort and validated the same classification effect as the public patients. All these findings help gain a deeper understanding about the pathogenesis of ccRCC and provide useful clues for optimizing therapeutic schemes based on the molecular subtype analysis.

INTRODUCTION

Renal cell carcinoma (RCC) is one of the ten most frequently diagnosed cancers worldwide.[1,2] Clear cell RCC (ccRCC) ranks the first common pathological subtype of RCC, accounting for approximately 75% of all RCC cases.[3] Although surgical resection is an effective therapeutic modality for early localized ccRCC, about 30% of patients with localized disease eventually develop metastasis.[3] The five-year disease-related survival (DRS) for metastatic ccRCC is only 12%. Molecular targeted therapies such as the use of tyrosine kinase inhibitors (TKIs) and immune checkpoint inhibitors (ICIs) are regarded as the standard front-line regimens for advanced and metastatic ccRCC.[4,5] However, the common low response rate induces drug resistance and disease progression in most patients due to the obvious individual response to treatment.[6] Therefore, distinguishing high-risk patients and choosing sensitive targeted therapies represent a growing challenge.

TNM staging and Fuhrman grade are generally accepted as common clinical indicators for assessing the progression of RCC. However, patients with the same clinical stage or grade may have different outcomes. Several studies have shown that molecular signatures can accurately predict the risk stratification of patients with ccRCC, including tumor recurrence, tumor-specific survival, and sensitivity to ICIs, and angiogenesis inhibitors.[7–9] Recent research has demonstrated that RCC tumor tissues can be categorized into seven molecular subsets by RNA transcriptomic analysis, which are associated with differential clinical outcomes to angiogenesis inhibitors or with an ICI.[9] However, most existing studies have mainly focused on the bulk tissue to classify molecular subtypes with prognostic value but ignored differences in the distinct single-cell type, much less in the number of cell populations associated with the efficacy of combined therapy. As a result, the properties and functions of key cell populations cannot be accurately obtained by bulk-seq study.

It is supposed that distinct RCC subtypes originate from different types of highly heterogeneous renal epithelial cells.[10] For example, ccRCC tumors used to be considered originating from proximal tubular cells,[11] and therefore, most research efforts were devoted to find potential cells of origin that share transcriptional characteristics with the ccRCC epithelium. However, with the booming of the cell stem cell (CSC) theory in recent years, they have been recognized as small subsets of cancer cells with self-renewal ability, which can stimulate tumor growth and lead to tumor heterogeneity.[12] CSCs also participate in tumor aggressiveness recurrence and metastasis, including resistance to treatment.[13] The biological characteristics of ccRCC cells may have appeared as different branches in the process of CSC differentiation, with different phenotypic states and distinct interactions with the microenvironment. As highly vascularized ccRCC exhibits high levels of immune cell infiltration,[14] metastatic ccRCC should be able to well respond to antiangiogenics and immunotherapies,[15,16] but reality is not desirable as

[1]Department of Urology, Xinhua Hospital, School of Medicine, Shanghai Jiaotong University, 1665 Kongjiang Road, Shanghai 200092, China
[2]Department of Urology, Third Affiliated Hospital of the Second Military Medical University, 700 Moyu North Road, Shanghai 201805, China
[3]Department of Pathology, Xinhua Hospital, School of Medicine, Shanghai Jiaotong University, 1665 Kongjiang Road, Shanghai 200092, China
[4]These authors contributed equally
[5]Lead contact
*Correspondence: gansishun20101111@163.com (S.-s.G.), brilliant212@163.com (W.Z.), cuixingang@xinhuamed.com.cn (X.-g.C.)
https://doi.org/10.1016/j.isci.2023.108370

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Table 1. Clinical information of ccRCC patients in scRNA-seq analysis

Patient ID	Gender	Age	Grade	Stage	Cell number
RCC1	Female	62	1	III	11824
RCC2	Male	46	2	I	6719
RCC3	Female	58	2	I	4530
RCC4	Male	65	3	I	9091
RCC5	Male	59	3	IV	5169
RCC6	Male	70	2	II	8711
RCC7	Female	71	2	II	13999
RCC8	Female	49	3	III	5775
RCC9	Male	65	3	IV	9381
RCC10	Female	79	2	IV	6591
RCC11	Female	69	2	II	7427

expected. Therefore, it deserves to investigate the differentiation branches of CSCs and their interactions with immune cells based on the molecular features, whose future may determine the tumor biological behavior and response to drug therapies.

Single-cell RNA sequencing (scRNA-seq) analysis can identify various cell types or physiological states and profile molecular characterization on thousands of cells, thereby revealing specific biological characteristics of various cell types.[17,18] Monocle3 algorithm in scRNA-seq can be used to reveal changes in cell-fate-dependent gene expression, splicing patterns, and allelic imbalances, which was potentially a method to subtyping cancer cells based on the differentiation of CSCs. In our previous study,[19] we used scRNA-seq analysis to characterize the complex cellular ecosystem and clinical biological evolution of the fatal collecting duct renal rCC (CDRCC) and predict potential intervention strategies for this deadly disease. Bi et al.[20] explored differences in the tumor microenvironment (TME) in metastatic RCC patients before or after ICI treatment by scRNA-seq analysis. However, these studies did not systematically establish RCC single-cell molecular subtypes to distinguish differential clinical outcomes.

Therefore, we used scRNA-seq analysis to identify unbiased cell subpopulations of ccRCC in 86,042 cells from patients with different stages and grades. The subpopulations of ccRCC cells that we identified were involved in several specific biological functions and associated with differential clinical outcomes. Three subtypes were identified to predict the prognostic risk of patients with ccRCC and assess the treatment response of the angiogenesis inhibitor and the checkpoint inhibitor. In addition, we tested the subtyping model through immunohistochemical (IHC) staining in our ccRCC tissue cohort (n = 286), which validated the same classification effect as the patients in public datasets. The biological and clinical significance of this study may shed light on accurate prediction and personalized treatment strategies for ccRCC.

RESULTS

Human ccRCC cells present significant changes as compared with normal cells

To understand the molecular characteristics of ccRCC, we collected 11 fresh ccRCC tumor tissues from patients with different stages and grades (Table 1) and digested them into single cells for quality screening and cell sequencing. As a result, we isolated 86,042 ccRCC cells for subsequent analysis by classifying them into 9 clusters for unsupervised clustering analysis based on the specific expression of cell-type signal genes, including cancer cells, distal tubule cells (DT), proximal tubule cells (PT), collecting duct cluster cells (CT), endothelial cells, fibroblasts, B cells, T cells, and myeloid cells (Figures 1A and S1A; Table S2). The uniform manifold approximation and projection (UMAP) plot showed the distribution of the cells in each sample (Figure 1B). Epithelial cells (both normal and malignant) constituted the main proportion of renal cells and cancer cells, accounting for about 60% of all cells (Figure S1B).

Each cell cluster was further subdivided into several subclusters using unsupervised clustering analysis. Cancer cells were subdivided into five cancer cell subclusters (ccRCC1-5 subclusters) (Figure 1C) and five cancer stem cell (CSC) subclusters (CSC and cancer stem-like cell [CSCL] 1–4 subclusters) (Figure 1D); T cells were subdivided into CD4+T, CD8+T (including CD8+T terminally exhausted [CD8+Tex], CD8+T terminally exhausted effector [CD8+Tex-eff], and CD8+T progenitor-like [CD8+T-pro]), nature killer T (NKT) cell, regulatory T cell (Treg), and follicular helper T cell (Tfh) (Figure 1E); B cells were subdivided into B cells and plasma cells (Figure 1F); myeloid cells were subdivided into tumor-associated macrophage (TAM), M1 macrophage (M1Φ), M1 macrophage (M2Φ), dendritic cell (DC), and monocyte (Figure 1G); endothelial cells were subdivided into endothelial1, endothelial2, and endothelial3 (Figure 1H). Thus, a total of 30 cell populations were identified in ccRCC tissues. Significant differences in gene expression were clearly exhibited between each subcluster, suggesting that cell populations were identified successfully (Figure 1H). Correlation analysis showed that cell subclusters in the same cell cluster were highly concentrated in the adjacent area, further confirming the reliability of the subclusters (Figure S1C). The distribution of unique molecular identifiers (UMIs) in each cell population is shown in Figure S1D. Each subcluster contained various cells from multiple samples, implying the rationality of the samples (Figure 1J).

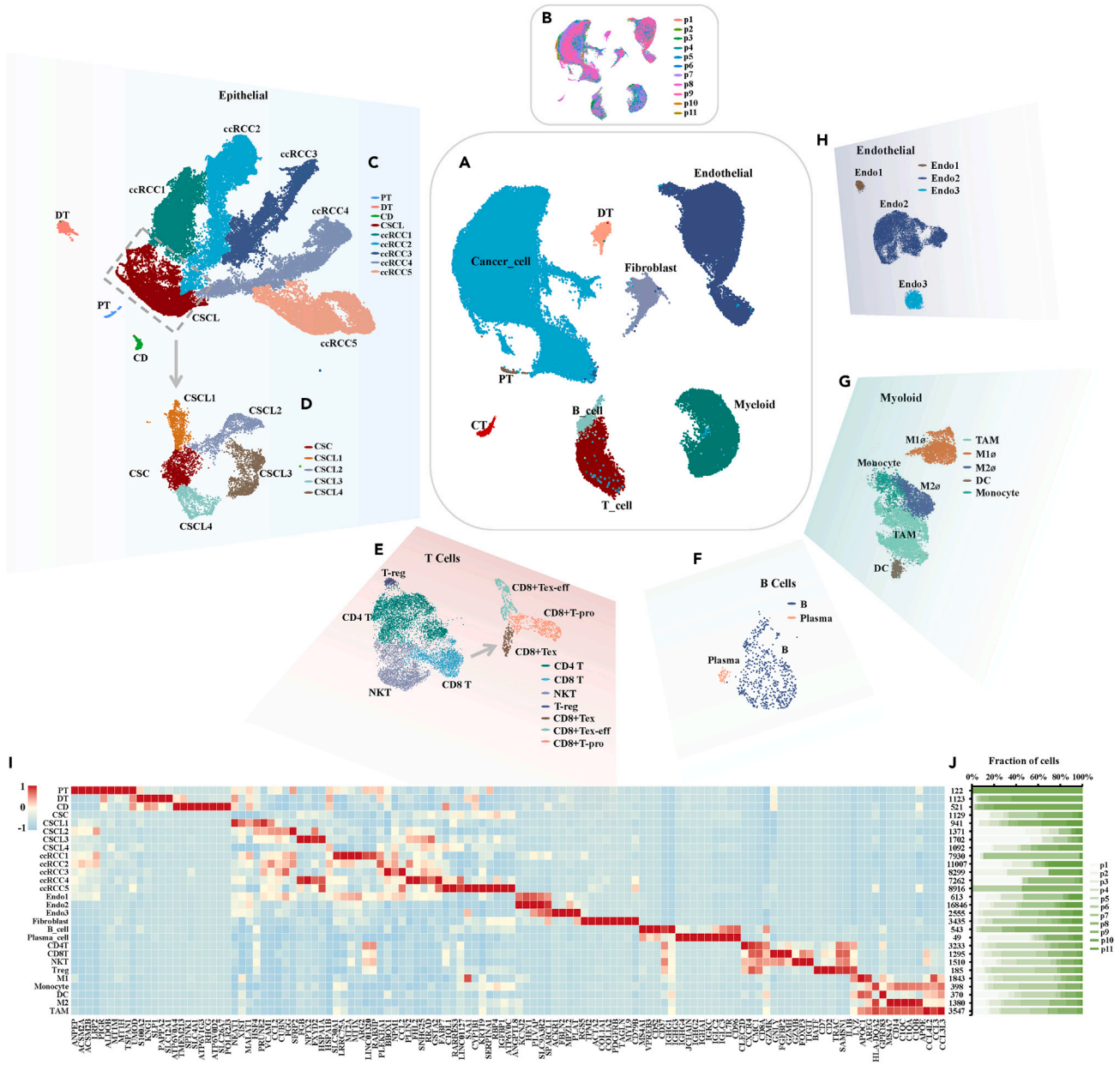


Figure 1. Identification of human ccRCC cell populations

(A) Unsupervised clustering for cells isolated from 11 ccRCC tissue samples. Nine major clusters were annotated.

(B) Cells in each sample (p1-p11) displayed by UMAP plot.

(C and D) Cancer cells was subdivided into five cancer cell subclusters (ccRCC1-5) and five cancer stem cell (CSC) subclusters (CSC and CSCL1-4).

(E) T cells were subdivided into CD4+T, CD8+T (including CD8+Tex, CD8+Tex-eff, and CD8+T-pro), NKT, T-reg, and Tfh.

(F) B cells were subdivided into B cells and plasma cells.

(G) Myeloid cells were subdivided into TAM, M1Φ, M2Φ, DC, and monocyte.

(H) Endothelial cells were subdivided into endothelial cells 1–3.

(I) Heatmap of gene expression in each subcluster.

(J) The bar plot of fraction of cells from multiple samples in each subcluster.

The distinct transcriptome program in epithelial cells suggests the potential tumor cell evolution origin

Epithelial cells are the main component of the ccRCC tumor tissue (Figure S1B) and the origin of ccRCC as well.[3] To distinguish the heterogeneity of epithelial cells, we innovatively applied the monocle3 method of cell differentiation to identify cancer cell subpopulations. As the monocle plot shown, cancer cell cluster was subdivided into six cancer-related subclusters, including one cancer stem-like cell (CSCL)

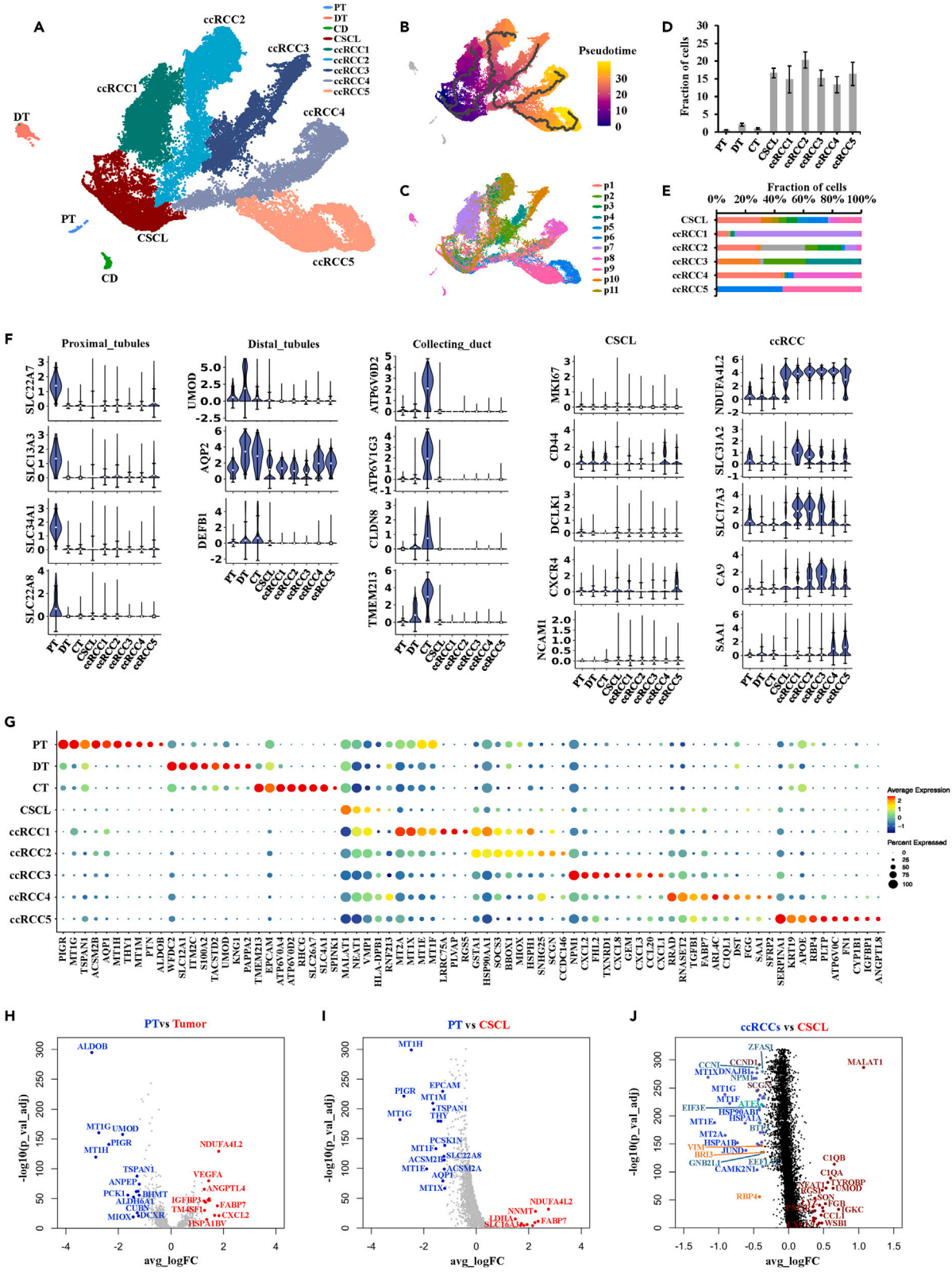

A
PT
DT
CD
CSCL
ccRCC1
ccRCC2
ccRCC3
ccRCC4
ccRCC5
B
Pseudotime
C
p1
p2
p3
p4
p5
p6
p7
p8
p9
p10
p11
D
Fraction of cells
E
Fraction of cells
0% 20% 40% 60% 80% 100%
F
Proximal_tubules
Distal_tubules
Collecting_duct
CSCL
ccRCC
SLC22A7
SLC13A3
SLC34A1
SLC22A8
UMOD
AQP2
DEFB1
ATP6V0D2
ATP6V1G3
CLDN8
TMEM213
MKI67
CD44
DCLK1
CXCR4
NCAM1
NDUFA4L2
SLC31A2
SLC17A3
CA9
SAA1
G
Average Expression
Percent Expressed
H
PTvs Tumor
avg_logFC
-log10(p_val_adj)
I
PT vs CSCL
J
ccRCCs vs CSCL

Figure 2. The distinct transcriptome program in epithelial cells

(A) Monocle3 analysis: cancer cell cluster was subdivided into six cancer-related subclusters, including CSCL and ccRCC 1–5.

(B) Pseudotime trajectory depicted CSCs cells (the start of differentiation process) can differentiate into ccRCC1-5 cells (the end of differentiation process). The legend showed the differentiation age (0–30).

(C) ccRCC cells in each sample (p1–p11) displayed by UMAP plot.

(D and E) The general distribution fraction of cancer cells among CSCL and ccRCC1-5 subclusters. Data are represented as mean ± SEM.

(F) Violin plot showed the well-known ccRCC-related cell markers to check the characteristics of ccRCC cell clusters. Data are represented as mean ± SEM.

(G) Dot plot expression of gene signature in each cancer cell subcluster and normal kidney cell subclusters.

(H–J) Volcano plot showed the gene characteristic differences between cancer cells and normal renal epithelial cells. (H) PT VS Tumor; (I) PT VS CSCL; (J) ccRCC1-5 VS CSCL. Red for upregulated genes; blue and other colors for downregulated genes.

subcluster and five branches with five cancer cell subclusters (ccRCC 1–5) (Figure 2A). According to monocle pseudotime trajectory and velocity analysis that we applied to depict the differentiation trajectory of cells, CSC cells, as the starting point of the differentiation process, could differentiate into ccRCC1-5 cells (Figures 2B and S2A). The general distribution patterns of cancer cells were similar between CSCL and ccRCC1-5 subclusters (Figure 2D). ccRCC-4 and ccRCC-5 subclusters, which were mainly derived from RCC 1, 6, 8, and 9, were at the relatively terminal stage of differentiation (Figures 2B, 2C, and 2E). ccRCC-1 subcluster was mainly derived from RCC 7, which was at the relatively early stage of differentiation (Figures 2B, 2C, and 2E).

To further verify whether the characteristics of ccRCC cell clusters that we annotated were correct, we used previously reported ccRCC-related cell markers to check the characteristics of ccRCC cell clusters. It was found that the known ccRCC CSC markers CD44,[19] CXCR4,[21] DCLK1,[22] NCAM1,[23] and MKI67[24] were highly expressed in CSCL cluster (Figure 2F). Compared with normal kidney cells, ccRCC cell clusters generally had higher expression of acknowledged ccRCC-related markers, such as CA9,[25] NDUFA4L2,[26] SAA1,[27] SLC17A3,[28] and SLC31A2[29] (Figure 2F). In addition, the markers for proximal tubules, distal tubules, and collecting duct were significantly highly expressed in PT, DT, and CD subclusters, respectively (Figure 2F). Moreover, as the heatmap shown, the expression of signature genes could clearly distinguish each cancer cell subcluster and normal kidney cell subclusters (DT, PT, and CD cells) (Figure 2G). Changes of partial gene expression during cancer cell differentiation trajectories are shown in Figure S2B. For example, VEGFA, T1MP, PLVAP, and HSPA1B, which are known to be associated with angiogenesis, were upregulated during CSCL-ccRCC1 differentiation trajectory. GATM and ECH1, which are known to regulate the biosynthesis of interleukins, were gradually increased in CSCL-ccRCC2 differentiation trajectory. The expression of FHL2, CXCL8, and TFPI, which are known to regulate the process of phosphorylation, showed an upward trend during CSCL-ccRCC3 differentiation trajectory. MDK and C1QL1, which are known to promote the process of cell proliferation, were gradually enhanced in CSCL-ccRCC4 differentiation trajectory. MYL9, SPON2, PLTP, and KRT19, which are known to regulate the process of immune response, were significantly upregulated during CSCL-ccRCC5 differentiation trajectory.

Knowing that ccRCC originates from renal tubular epithelial cells, we focused on differences in gene characteristics between cancer cells and normal renal epithelial cells. As shown in the volcano map, compared with normal epithelial cells, cancer cells were enriched in hypoxia-responsive and angiogenic genes, such as NDUFA4L2,[30] ANGPTL4,[31] IGFBP3,[32] and VEGFA[33] (Figure 2H), and CSCL cells were enriched in genes related to self-renewal, uncontrolled proliferation, and differentiation, such as NNMT,[34] LDHA,[35] and SLC16A3[36] (Figure 2I). Similarly, volcano map from "ccRCC versus CSCL" showed that specific genes were enriched separately in ccRCC cell clusters and CSCL cell cluster (Figure 2J). Comparison of gene expressions and transcription factor (TF) activity between normal epithelial cells (PT, CD, and DT cells) are shown in Figures S3A–S3E.

Specific molecular features of the five differentiation branches of ccRCC cells

To further identify the biological features of the five branches of ccRCC cells, gene set variation analysis (GSVA) was applied in the ccRCC subclusters. The enrichment of gene ontology (GO) pathway in CSCL subcluster was similar with that in ccRCC1 subcluster (Figure 3A), indicating that CSCL and ccRCC1 cells were the primitive stage of ccRCC cell differentiation, as the lower GO enrichment scores pointed to the undeveloped biological functions. In addition, the enrichment score of GO pathway was significantly increased in ccRCC 2–5 subclusters compared with CSCL and ccRCC1 subclusters (Figure 3A). The cell-cycle stage distribution of cancer cells showed that CSCL cells were significantly enriched in G2/M and S stage (Figure 3B), which is associated with the proliferation, pluripotency, and differentiation of stem cells.

To investigate genetic heterogeneity of large-scale CNV in ccRCC subclusters, inferCNV was calculated based on scRNA data. As inferCNV analysis showed (Figure 3C), chromosomal loss in 3p was commonly observed in ccRCC cells (CSCL and ccRCC 1–5 cells), and this phenomenon was not observable in normal kidney cells. 3p loss is nearly a universal event in ccRCC,[3] which further implies the accuracy of the ccRCC cells that we identified. Other common ccRCC CNVs[37] were also observed in our study, including 5q gain in ccRCC 3–4 cells, 14q loss in ccRCC 2–5 cells, 7 gain in ccRCC 3–5 cells, and 9 loss in ccRCC 4–5 cells (Figure 3C). Interestingly, chromosomal loss in 9p and gain in 20p were only observed in ccRCC 4 and ccRCC 5 cells (Figure 3C), which have also been reported by previous studies, which observed that these variation events were associated with poor prognosis of ccRCC patients.[38,39]

We further analyzed the enrichment of the specific pathways in ccRCC1-5 subclusters to define their biological characteristics and found that ccRCC1 subclusters were more enriched in angiogenesis-related pathways than that in the other subcluster. ccRCC2 subcluster showed the enrichment of interleukin-biosynthesis-related pathways; ccRCC3 subcluster showed the enrichment of phosphorylation-activity-related pathways; ccRCC4 subcluster showed the enrichment of cell-proliferation-related pathways; and ccRCC5 subcluster showed the enrichment of immunomodulatory signaling pathways (Figure 3D).

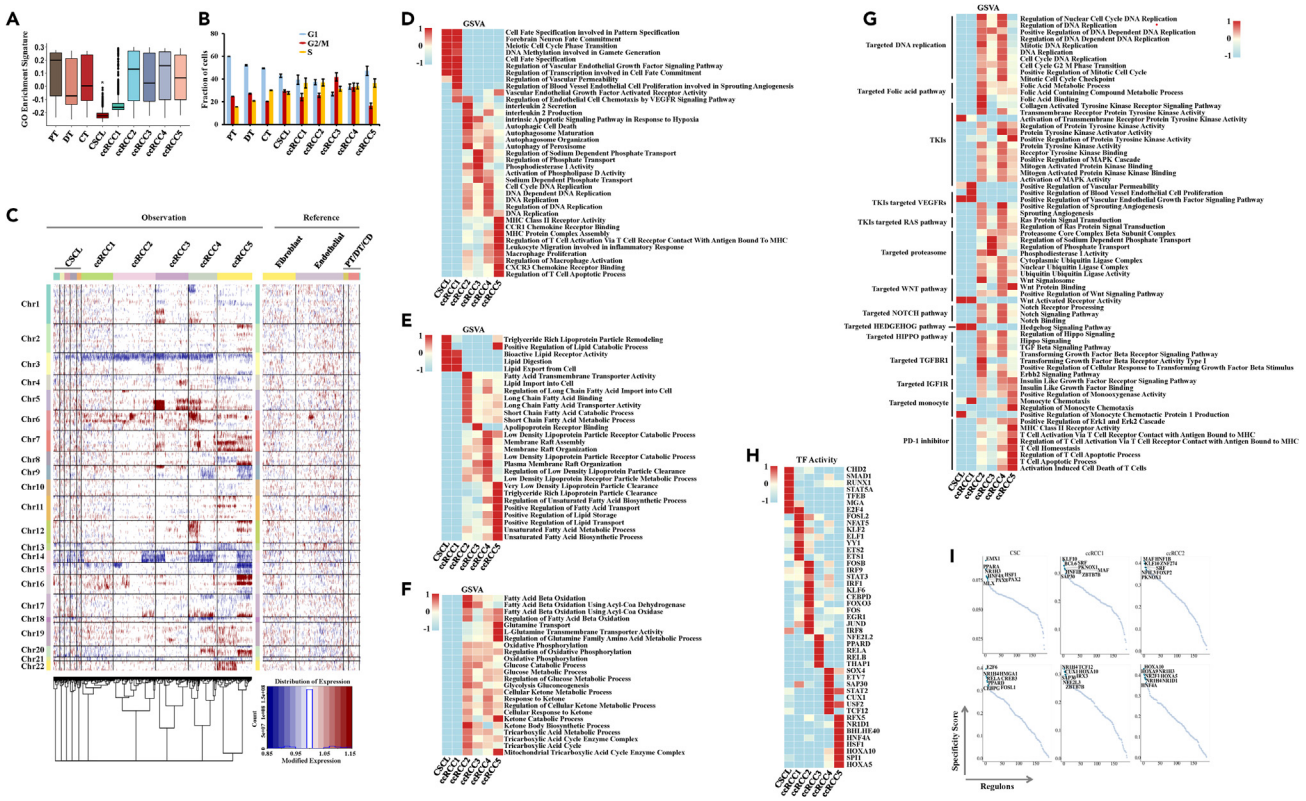


Figure 3. Specific molecular features of the five differentiation branches of ccRCC cells

(A) The GO enrichment signature score of cancer cell subclusters and normal kidney cell subclusters. Data are represented as mean ± SEM.

(B) The fraction of cancer cells in each cell-cycle stage. Blue for G1 stage; red for G2/M stage; yellow for S stage. Data are represented as mean ± SEM.

(C) The inferCNV heatmap for each ccRCC group. Red for gain; blue for loss. The reference cells were normal tubular cell, endothelial, and fibroblasts.

(D) The heatmap for activated cancer-related pathways in each cancer cell group. Red for upregulation; blue for downregulation.

(E and F) Enrichment analysis heatmap of the lipid-metabolism-related pathways in each cancer cell group.

(G) Heatmap for pathways related to targeted therapies among the six ccRCC subpopulations.

(H) Heatmap for transcriptional factor activations in each cancer cell group.

(I) Regulons activity score in CSC and ccRCC 1–5.

ccRCC is characterized by lipid metabolism dysfunction.[40] Abnormal lipid metabolism in ccRCC promotes cellular lipid accumulation, which provides the additional ATP for enhancing the viability of cancer cells to promote disease progression.[41,42] We further analyzed the enrichment of lipid metabolism pathways in ccRCC1-5 subclusters. As shown in the enrichment heatmap of the pathways, each cancer cell subcluster presented up-regulation of meaningful pathways that specifically regulated lipid metabolism (Figure 3E). Interestingly, ccRCC5 subcluster accelerated the accumulation of lipids through biosynthesis, storage, and transportation (Figure 3E). The abundance of metabolites of the glutaminolytic pathway and ketone catabolic process was higher in ccRCC5 subcluster (Figure 3F). However, the fatty acid oxidation (FAO)-related pathways were enriched in ccRCC2 and ccRCC3 subclusters (Figures 3E and 3F). Abnormal lipid accumulation and glutamine metabolism have been found to be critical in promoting the progression of ccRCC,[41,43] but upregulation of FAO in cancer cells is equivalent to tumor cell "slimming," which represses the progression of ccRCC.[41] Therefore, cancer cells in ccRCC5 subcluster may be a higher prognostic risk compared with cancer cells in other subclusters, especially in ccRCC2 and ccRCC3 subclusters.

In addition, we analyzed the pathways related to targeted therapies among the five ccRCC subpopulations (Figure 3G). For example, the targeted therapy-related pathway of ccRCC1 subcluster was enriched in membrane receptor activation of tyrosine kinases such as VEGFR, suggesting that ccRCC1 subcluster may be sensitive to sunitinib. Interestingly, ccRCC4 and ccRCC5 subclusters were not only enriched in T cell activation, T cell homeostasis, and MHC class II receptor activity but also upregulated TKI targeted VEGFR pathways, including positive regulation of sprouting angiogenesis, sprouting angiogenesis, and vascular-endothelial-growth-factor-activated receptor activity (Figures 3D and 3G). These findings suggest that ccRCC4 and ccRCC5 subclusters may be sensitive to TKI-targeted therapy combined with ICI immunotherapy.

To further study the transcriptional program in the cancer cell populations, transcriptional factor (TF) analysis was used. The result highlighted that different cancer cell subclusters had different TF activations (Figures 3H and 3I). FOSL2, NFAT5, KLF2, ELF1, YY1, ETS1, and ETS2, which are known to promote tumor angiogenesis,[44–48] were generally up-regulated in the ccRCC1 subcluster. TCF12, SOX4, STAT2, and

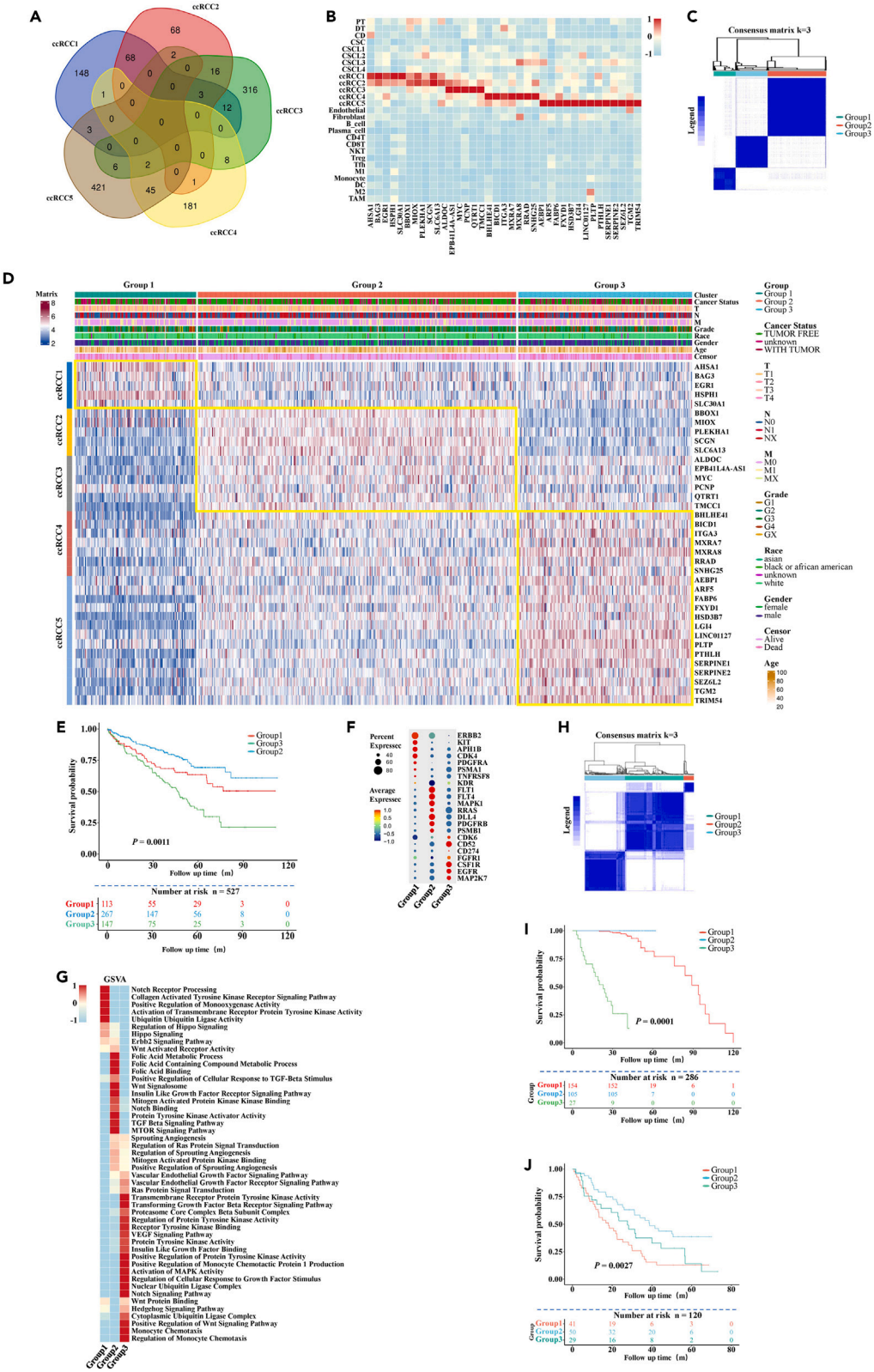

A
ccRCC1
ccRCC2
ccRCC3
ccRCC4
ccRCC5
B
C
Consensus matrix k=3
Legend
Group1
Group2
Group3
D
Group 1
Group 2
Group 3
Matrix
Cluster
Cancer Status
T
N
M
Grade
Race
Gender
Age
Censor
Group
Cancer Status
TUMOR FREE
unknown
WITH TUMOR
Grade
Race
asian
black or african american
unknown
white
Gender
female
male
Censor
Alive
Dead
E
Survival probability
P = 0.0011
Follow up time (m)
Number at risk n = 527
Group1 113 55 29 3 0
Group2 267 147 56 8 0
Group3 147 75 25 3 0
F
Percent Expressec
Average Expressec
ERBB2
KIT
APH1B
CDK4
PDGFRA
PSMA1
TNFRSF8
KDR
FLT1
FLT4
MAPK1
RRAS
DLL4
PDGFRB
PSMB1
CDK6
CD52
CD274
FGFR1
CSF1R
EGFR
MAP2K7
G
GSVA
Notch Receptor Processing
Collagen Activated Tyrosine Kinase Receptor Signaling Pathway
Positive Regulation of Monooxygenase Activity
Activation of Transmembrane Receptor Protein Tyrosine Kinase Activity
Ubiquitin Ubiquitin Ligase Activity
Regulation of Hippo Signaling
Hippo Signaling
Erbb2 Signaling Pathway
Wnt Activated Receptor Activity
Folic Acid Metabolic Process
Folic Acid Containing Compound Metabolic Process
Folic Acid Binding
Positive Regulation of Cellular Response to TGF-Beta Stimulus
Wnt Signalosome
Insulin Like Growth Factor Receptor Signaling Pathway
Mitogen Activated Protein Kinase Kinase Binding
Notch Binding
Protein Tyrosine Kinase Activator Activity
TGF Beta Signaling Pathway
MTOR Signaling Pathway
Sprouting Angiogenesis
Regulation of Ras Protein Signal Transduction
Regulation of Sprouting Angiogenesis
Mitogen Activated Protein Kinase Binding
Positive Regulation of Sprouting Angiogenesis
Vascular Endothelial Growth Factor Signaling Pathway
Vascular Endothelial Growth Factor Receptor Signaling Pathway
Ras Protein Signal Transduction
Transmembrane Receptor Protein Tyrosine Kinase Activity
Transforming Growth Factor Beta Receptor Signaling Pathway
Proteasome Core Complex Beta Subunit Complex
Regulation of Protein Tyrosine Kinase Activity
Receptor Tyrosine Kinase Binding
VEGF Signaling Pathway
Protein Tyrosine Kinase Activity
Insulin Like Growth Factor Binding
Positive Regulation of Protein Tyrosine Kinase Activity
Positive Regulation of Monocyte Chemotactic Protein 1 Production
Activation of MAPK Activity
Regulation of Cellular Response to Growth Factor Stimulus
Nuclear Ubiquitin Ligase Complex
Notch Signaling Pathway
Wnt Protein Binding
Hedgehog Signaling Pathway
Cytoplasmic Ubiquitin Ligase Complex
Positive Regulation of Wnt Signaling Pathway
Monocyte Chemotaxis
Regulation of Monocyte Chemotaxis
H
Consensus matrix k=3
I
P = 0.0001
Number at risk n = 286
J
P = 0.0027
Number at risk n = 120

Figure 4. ccRCC subsets in scRNA-seq were mapped to the three molecular subtypes and associated with clinical prognosis

(A) The Venn plot showed the intersections between the marker genes of ccRCC1-5 subclusters.

(B) Heatmap for expression of 35 most specific marker genes in cancer cell populations.

(C) Optimized discrimination effect of the 527 TCGA-KIRC samples in the three subtypes based on the tumor profiles with 5 gene panels (35 genes) of ccRCC1-5 subclusters.

(D) Complex heatmap for gene expression clustering. ccRCC1 subcluster enriched in group 1, ccRCC2 and ccRCC3 subclusters enriched in group 2, and ccRCC4 and ccRCC5 subclusters enriched in group 3.

(E) Survival analysis for group 1–3 in TCGA. p = 0.0011.

(F) Heatmap for the gene expression related to tyrosine kinase.

(G) Heatmap for GSVA analysis of VEGF or angiogenesis-related signaling pathway in group 1–3.

(H) Consensus clustering for our own ccRCC patients (n = 286). The optimized discrimination effect of three groups in our cohort.

(I) Survival analysis for group 1–3 in our cohort. p = 0.0001.

(J) Survival analysis for group 1–3 in patients received ICIs therapy from Checkmate025. p = 0.0027.

CUX1, which are known to increase the risk of tumor proliferation and metastasis,[49–52] were activated transcriptionally in the ccRCC4 subcluster.

Molecular classification is correlated with prognostic risk stratification and different clinical outcomes to targeted therapies

GSVA, CNV, and TF activation presented specific characteristics in the cancer cell populations, thus we tried to explore the molecular classification based on ccRCC1-5 subclusters. The intersections between the marker genes of ccRCC1-5 subclusters were quite limited (Figure 4A), implying that the marker genes of ccRCC1-5 subclusters could be used to build scRNA-seq-based classifier. We selected 35 most specific marker genes of cancer cell populations, including 5 markers (AHSA1, BAG3, EGR1, HSPH1, and BBOX1) in ccRCC1 subcluster, 5 markers (MIOX, PLEKHA1, SCGN, SLC6A13, and ALDOC) in ccRCC2 subcluster, 5 markers (EPB41L4A-AS1, MYC, PCNP, QTRT1, and TMCC1) in ccRCC3 subcluster, 7 markers (BHLHE41, BICD1, ITGA3, MXRA7, MXRA8, RRAD, and SNHG25) in ccRCC4 subcluster, and 13 markers (AEBP1, ARBP1, ARF5, FABP6, FXYD1, HSD3B7, LGI4, LINC01127, PLTP, PLTP, PYHLH, SERPINE1, SERPINE2, SEZ6L2, TGM2, and TRIM54) in ccRCC5 subcluster (Figure 4B).

To explore the possible application of cancer cell population marker genes in ccRCC molecular classification, tumor profiles of 5 randomly selected gene panels (35 genes) of ccRCC1-5 subclusters were used to classify the 527 samples in TCGA dataset into subtypes. The result showed that the three subtypes achieved an optimized discrimination effect of the ccRCC samples (Figure 4C). The markers of ccRCC1 subcluster were significantly enriched in group 1, the markers of ccRCC2 and ccRCC3 subclusters were observed in group 2, and the markers of ccRCC4 and ccRCC5 subclusters were increased in group 3 (Figure 4D). We subsequently evaluated the ccRCC prognostic risk in each group. Patients in group 3 (ccRCC4 and ccRCC5 subclusters) were demonstrated to have a worse survival outcome compared with patients in group 1 (ccRCC1 subclusters) and group 2 (ccRCC2 and ccRCC3 subclusters) (Figure 4E). We validated the model internally in the validation cohort from our own ccRCC patients (n = 286). As expected, the differences in prognostic risk between the three groups were consistent with the aforementioned results (Figures 4H and 4I).

Group 1 and group 3 clusters were up-regulated in angiogenesis-related pathways, such as activation of transmembrane receptor protein tyrosine kinase activity, VEGF signaling pathway, regulation of protein tyrosine kinase activity, and regulation of angiogenesis (Figures 4F, 4G, and S4). Group 2 cluster was enriched in energy-metabolism-related signaling pathways, and patients in this group exhibited the best survival outcomes, implying that tumor cells progressed slowly and did not require any treatment (Figures 4F, 4G, and S5). Of great interest, pathways related to ICI therapies were associated with increases in group 3 cluster, including regulation of the immune effector process and acute inflammatory response (Figures 4F, 4G, and S6). We further characterized the prevalence of the top altered genes in each group, and the prevalence of BAP1, PTEN, PBRM1, and KDM5C mutations was the highest in group 3 (inflammation/angiogenesis cluster), which is similar to previous research[9] (Figure S7). Therefore, we termed group 1 subcluster as angiogenesis activity, group 2 subcluster as indolent progress, and group 3 subcluster as inflammation+ angiogenesis activity.

We used patients (n = 120) with ICI therapy from checkmate025 dataset to verify the therapeutic efficacy of ICI therapy. As expected, we found that patients with ICIs in group 3 exhibited markedly better survival outcomes than patients in group 1, suggesting that patients in group 3 are more suitable for ICI treatment (Figure 4J). The disease control rate (DCR) of patients with ICI therapy was 79% (group 3), 68% (group 1), and 65% (group 2) (Figure S8A). Compared with untreated ccRCC, tumor purity was decreased and immune scores were increased significantly in ccRCC tissues with ICIs therapy in group 3 (Figure S8B), demonstrating that our molecular classification based on scRNA-seq analysis could be effectively applied to risk prognostication for ccRCC patients and provide useful information for personalized treatment.

ccRCC stem cells are associated with the development and progression of ccRCC

Owing to the vital role of CSCs in the development and progression of ccRCC,[53] we further analyzed the biological properties of CSCL cells in ccRCC tissues and found that CSCL cells could be subdivided into five subclusters by monocle analysis, including CSC subcluster, CSCL1 subcluster, CSCL2 subcluster, CSCL3 subcluster, and CSCL4 subcluster (Figure 5A). Pseudotime trajectory analysis revealed that CSC cells in CSC subcluster were differentiated into four branches corresponding to the cells of CSCL1-4 subclusters (Figure 5B). Cells in each CSC cell subcluster were derived from multiple samples (Figure 5C) and had a similar cell number of about 1,500 (Figure 5D). GO pathway

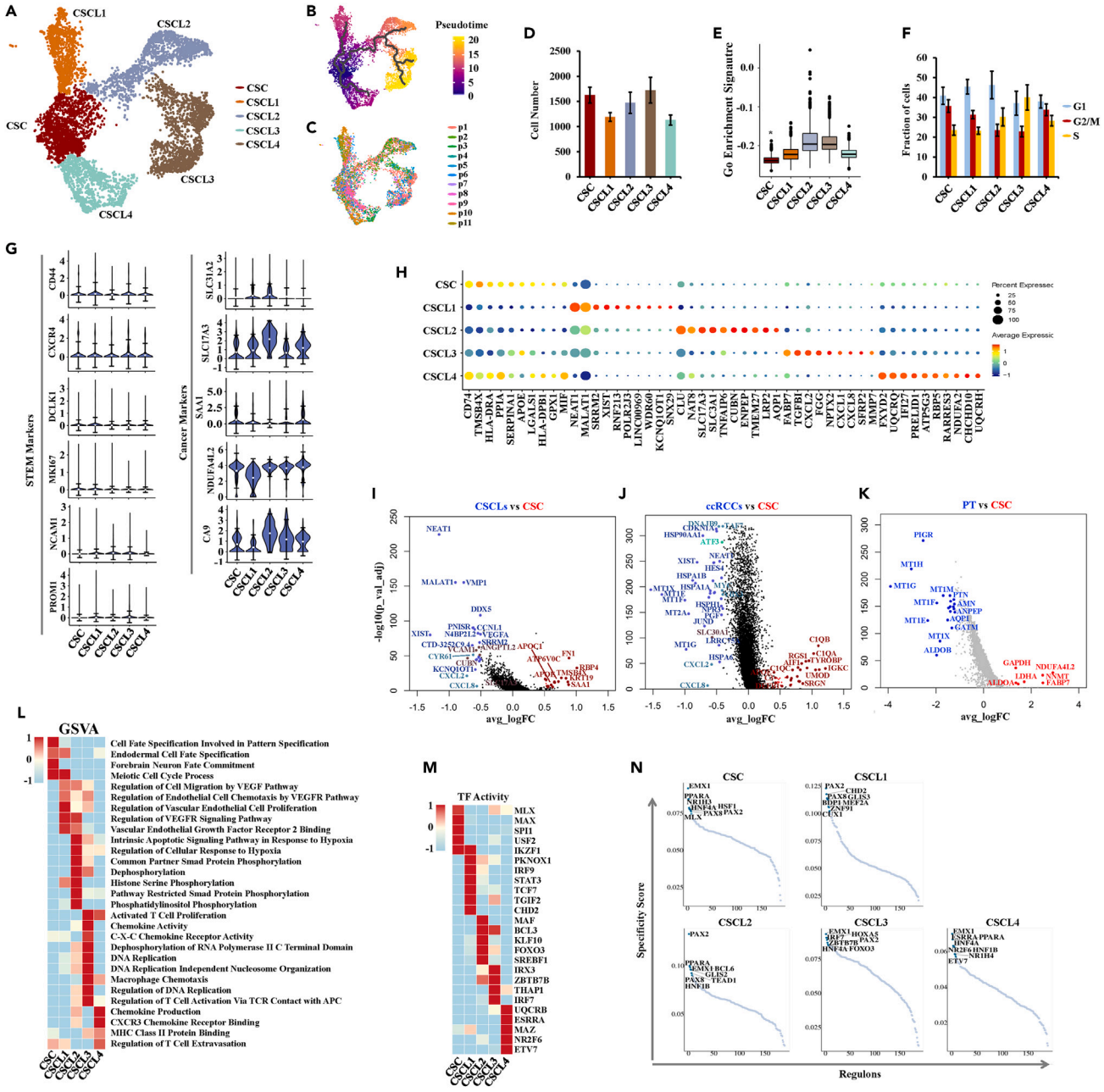


Figure 5. Identification of ccRCC stem cells

(A) UMAP plot of Monocle3 analysis showed that CSC cells could be subdivided into four CSCLs.

(B) Pseudotime trajectory depicted CSCs cells (the start of differentiation process) can differentiate into CSCLs 1–4 (the end of differentiation process). The legend showed the differentiation age (0–20).

(C) CSCs and CSCLs in each sample (p1–p11) displayed by UMAP plot.

(D)The cell number in CSC and CSCL1-4. Data are represented as mean ± SEM.

(E) The GO enrichment signature score of CSC and CSCL1–4. Data are represented as mean ± SEM.

(F) The fraction of CSC and CSCL1–4 in each cell-cycle stage. Blue for G1 stage; red for G2/M stage; yellow for S stage. Data are represented as mean ± SEM.

(G) Violin plot showed the well-known stem and cancer markers to check the characteristics of CSC and CSCL1–4. Data are represented as mean ± SEM.

(H) Dot plot expression of gene signature among CSC and CSCL1–4.

(I–K) Volcano plot showed the gene characteristic differences among CSCs, cancer cells, and normal renal epithelial cells. (I) CSCLs VS CSC; (J) ccRCC VS CSC; (K) PT VS CSC. Red for upregulated genes; blue and other colors for downregulated genes.

(L) The heatmap for activated cancer-related pathways in CSC and CSCL1–4. Red for upregulation; blue for downregulation.

(M) Heatmap for transcriptional factor activations in CSC and CSCL1–4.

(N) Regulons activity score in CSC and CSCL1–4.

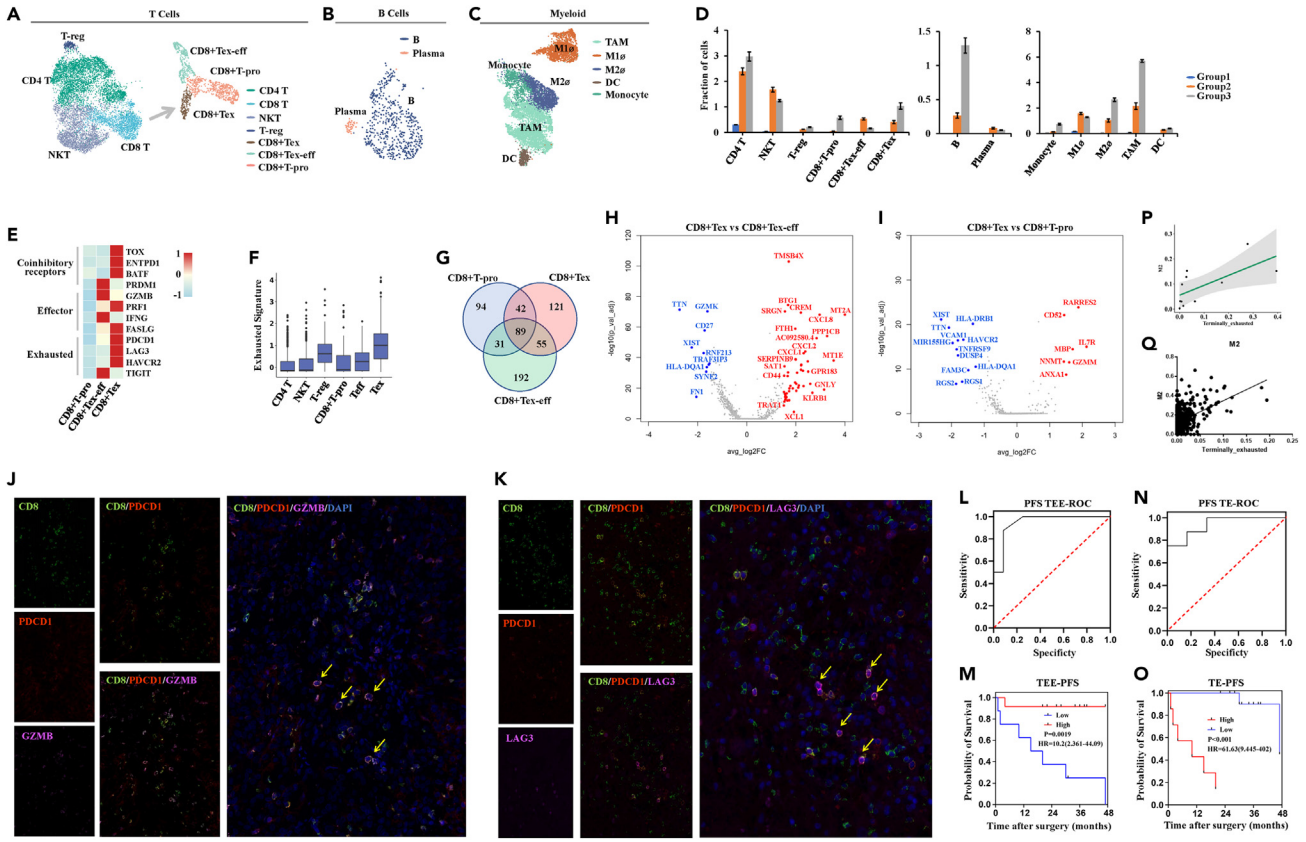


Figure 6. Identification of the immune microenvironment in ccRCC

(A–C) (Left to right) The subclusters of T cells, B cells, and myeloid cells by monocle3 analysis.

(D) The number of subclusters of T cells, B cells, and myeloid cells. Data are represented as mean ± SEM.

(E and F) Expression of exhaustion markers and signature in CD8+T-pro/CD8+Tex and CD8+T-pro/CD8+Tex-eff. Data are represented as mean ± SEM.

(G) Venn diagram of the signature genes among CD8+T-pro/CD8+Tex and CD8+T-pro/CD8+Tex-eff.

(H and I) Volcano plot showed the gene characteristic differences among CD8+T-pro/CD8+Tex and CD8+T-pro/CD8+Tex-eff. Red for upregulated genes; blue and other colors for downregulated genes.

(J and K) The IHC for CD8+Tex-eff [(J) CD8/PDCD1/GZMB] and CD8+Tex [(K) CD8/PDCD1/LAG3].

(L and M) The ROC and KM curve of PFS in patients with ICIs treatment based on IHC score of CD8+Tex-eff markers.

(N and O) The ROC and KM curve of PFS in patients with ICIs treatment based on IHC score of CD8+Tex markers.

(P and Q) M2Φ cells were positively associated with CD8+Tex cells.

enrichment was relatively high in CSCL2 and CSCL3 subcluster (Figure 5E). As mentioned in the previous cell-cycle analysis,[54] both the total CSC cells and CSC cells of CSC cell subcluster exhibited G1 phase arrest, indicating a better consistency of our data (Figure 5F).

All CSC cell subclusters highly expressed known stem cell markers, including CD44,[19] CXCR4,[21] DCLK1,[22] NCAM1,[23] MKI67,[24] and PROM1[55] (Figure 5G). In addition, known ccRCC markers (CA9,[25] NDUFA4L2,[26] SAA1,[27] SLC17A3,[28] and SLC31A2[29]) were commonly increased in CSC cell subclusters (Figure 5G). These results illuminate that identification of ccRCC CSCs was accurate. There were significant differences in the expression of signature genes between CSC cell subclusters (Figure 5H). Compared with CSCLs, ccRCCs, and PT cells, CSCs presented the characteristics of self-renewal and differentiation with high expression of stem-cell-related genes such as RBP4,[56] ANGPTL 2,[57] VCAM1,[58] AIF1,[59] NDUFA4L2,[60] and LDHA[35] (Figures 5I–5K).

GSVA analysis showed enrichment of CSCL1 subcluster in VEGFR signaling pathway and enrichment of CSCL4 subcluster in VEGFR signaling pathway (Figure 5L). TF analysis was performed to detect specific TF activation in CSC cell subclusters, and the result showed that SPI1,[61] USF2,[62] and HNF4A,[63] key TF regulating stem and progenitor cells in CSC subcluster, were significantly up-regulated as compared with CSCL subcluster (Figures 5M and 5N).

Immunosuppressive TME exerted by ccRCC is correlated with unfavorable prognosis

To explore the underlying mechanisms of response to immunotherapy in group 3 patients, further efforts are needed to gain further insights into the immune microenvironment. As shown in Figures 6A–6C and S9, T cells, B cells and myeloid cells were subdivided into several

subclusters. It is noteworthy that CD8 T cells were subdivided into three subclusters (CD8+Tex, CD8+Tex-eff, and CD8+T-pro) by monocle analysis. As expected, the number of CD8+Tex cells, M2Φ cells, and TAM cells in group 3 was significantly higher than that in the other groups, suggesting that ccRCC patients with characteristics of group 3 may present a worse prognosis (Figure 6D). With the pathological grade of ccRCC increasing, the number of CD8+Tex cells, M2Φ cells, and TAM cells also increased. However, radiology was unable to show the specific distribution of tumor stages of these cells (Figure S10). These findings of our study indicate that the terminal state of CD8 T cells presented two differentiation branches, including CD8+T-pro/CD8+Tex and CD8+T-pro/CD8+Tex-eff, rather than a linear differentiation model as the previous study.[64] CD8+T-pro cells, as relatively naive CD8 T cells, presented low expression of the specific marker genes of effector or exhausted state (Figures 6E and S11A). As for terminal $CD8^+$ T cells, exhaustion markers were upregulated in both CD8+Tex cells and CD8+Tex-eff cells (PDCD1, LAG3, HAVCR2 (encoding TIM-3), and ITGA4 for CD8+Tex cells; TIGIT and PDCD1 for CD8+Tex-eff cells) (Figures 6E, 6F, and S11A). Of effector molecules, we found that GZMB, IFNG, and TRGC2 were significantly upregulated in CD8+Tex-eff cells (Figures 6E and S11A). Although the expression of PRF1 and FASLG was also up-regulated in CD8+Tex cells (Figures 6E and S11A), only CD8+Tex-eff cells upregulated a host of chemokines (CXCL1, CXCL2, CXCL8, XCL1, and XCL2) (Figures 6H and S11A). As shown in Venn diagram analysis of the signature genes of three CD8 T subclusters, they shared functional similarities on one hand but on the other hand they also showed important differences (Figures 6G–6I). Therefore, we identified CD8+Tex cells as "dysfunctional terminal state" and CD8+Tex-eff as "functional effector terminal state."

Immunofluorescence analysis of our ccRCC tissues showed that CD8/PDCD1/LAG3 (markers for CD8+Tex cells) and CD8/PDCD1/GZMB (markers for CD8+Tex-eff cells) were co-expressed in some CD8 T cells (Figures 6J and 6K). Using PFS as the primary outcome, ROC curve showed that the best IHC score cutoff value for CD8+Tex-eff was 107.5 with an AUC of 0. 9479 (Figure 6L), and the survival analysis showed that high expression of CD8+Tex-eff indicated a better PFS (Figure 6M). However, the best IHC-score cutoff value for CD8+Tex was 97 with an AUC of 0.9375 (Figure 6N), and the survival analysis showed that a high expression of CD8+Tex markers indicated a worse PFS (Figure 6O). Compared with the untreated population, ccRCC patients with anti-PD-1 therapy in group 3 presented an interesting phenomenon that the number of CD8+Tex-eff cells was elevated and the number of CD8+Tex cells was decreased (Figure S11B). Thus, immune microenvironment reprogramming may be a potential cause of the response to ICI treatment in group 3 patients.

Ligand receptor mediates intercellular interactions in the ccRCC immune microenvironment

To better understand the interconnections between cells in the TME, cell-phone DB analysis was performed to detect the interactions between cell populations (Figure S12A). It was found that M2Φ cells were positively associated with CD8+Tex cells (Figures 6P and 6Q). Compared with CD8+Tex-eff cells, M2Φ/TAM cells maintained a tight association with CD8+Tex cells via the LGALS9/HAVCR2 receptor ligand complex (Figures 7A and 7C) by suppressing the cytotoxic viability of CD8+Tex cells via LGALS9-HAVCR2 (encoded by TIM3) interaction. Meanwhile, CD8+Tex cells were inferred to signal to M1Φ cells expressing CD55 via ADGRE5 (encoded by CD97), an interaction that may promote the polarization of M1 cells (Figures 7A and 7C). In addition, M2Φ/TAM cells enhanced the stability and function of CD8+Tex-eff cells and T-reg cells via LGALS9-CD44 interactions (Figures 7A and 7C).

We additionally observed cell-cell interactions between cancer cells and other cells (Figures 7B, 7C, and S12B). As expected, ccRCC1 cells acted on KDR1 and FLT1 receptors of endothelial cells by secreting VEGFA (Figures 7B, 7C, and S12B), thus promoting angiogenesis of endothelial cells in the TME.[65] Likewise, both cancer 4 and cancer 5 cells (group 3) secreted MDK and ANXA1, signaling to LRP1/SORL1 and FPR1/3 expressed on M2Φ/TAM cells (Figures 7B and 7C), interactions that induce chemotaxis of TAM cells to interfere with the biological function of CD8 T cells.[66,67] Furthermore, cancer 4 and cancer 5 cells were inferred to engage C3AR1 expressed on M2Φ/TAM cells via complement C3 (Figures 7B and 7C). Signaling through C3AR1 promotes the polarization of M2Φ/TAM cells to inhibit the cytotoxic viability of CD8+Tex cells via LGALS9-HAVCR2 interaction.[68]

DISCUSSION

Renal cell carcinoma is a highly heterogeneous malignant tumor. The same pathological type of RCC contains the presence of genetically different subpopulations, which is the possible cause of treatment failure and development of drug resistance.[69] Therefore, dissecting the molecular characterization of RCC with different subpopulations can help gain a better understanding of this disease. Based on scRNA-seq and the tumor differentiation view, our study provides insights into diverse biological phenotypes that are associated with the progression and therapeutic outcomes of ccRCC, which may help identify targets for future therapeutic development.

Some studies have attempted to stratify ccRCC patients based on the genomic and transcriptomic analysis of bulk tissue samples,[9,37] but the complex TME may confound the heterogeneity of tumor cells. scRNA-seq analysis seeks intra-tumor heterogeneity and cell-to-cell interconnections by splitting a mixed tumor into individual tumor cells, immune cells, and stromal cells. CSCs, as the origin of tumor cells, can differentiate into tumor cells of different biological characteristics.[70,71] Based on this point, we innovatively used the differentiation trajectories of CSCs to classify tumor cell subsets. We defined five terminal differentiation trajectories of ccRCC cells. It is worth noting that ccRCC2 and ccRCC3 cells originate from the close differentiation source, and a similar phenomenon was observed in ccRCC4 and ccRCC5 cells. As might be anticipated, ccRCC2 and ccRCC3 subclusters were automatically archived as a group (group 2) under unsupervised conditions in the molecular classification, as were ccRCC4 and ccRCC5 subclusters. We found that ccRCC patients in group 2 belonged to the low-risk cohort with better survival prognosis and better response to ICI therapy. Interestingly, group 3 patients belonged to the high-risk cohort and showed better responses to ICIS treatment compared with group 2 patients, an intermediate-risk cohort. Some cancer subtypes become highly pathogenic and resistant to drugs, potentially also due to changes in epigenetic levels, such as the influence of TFs. In our study, each ccRCC

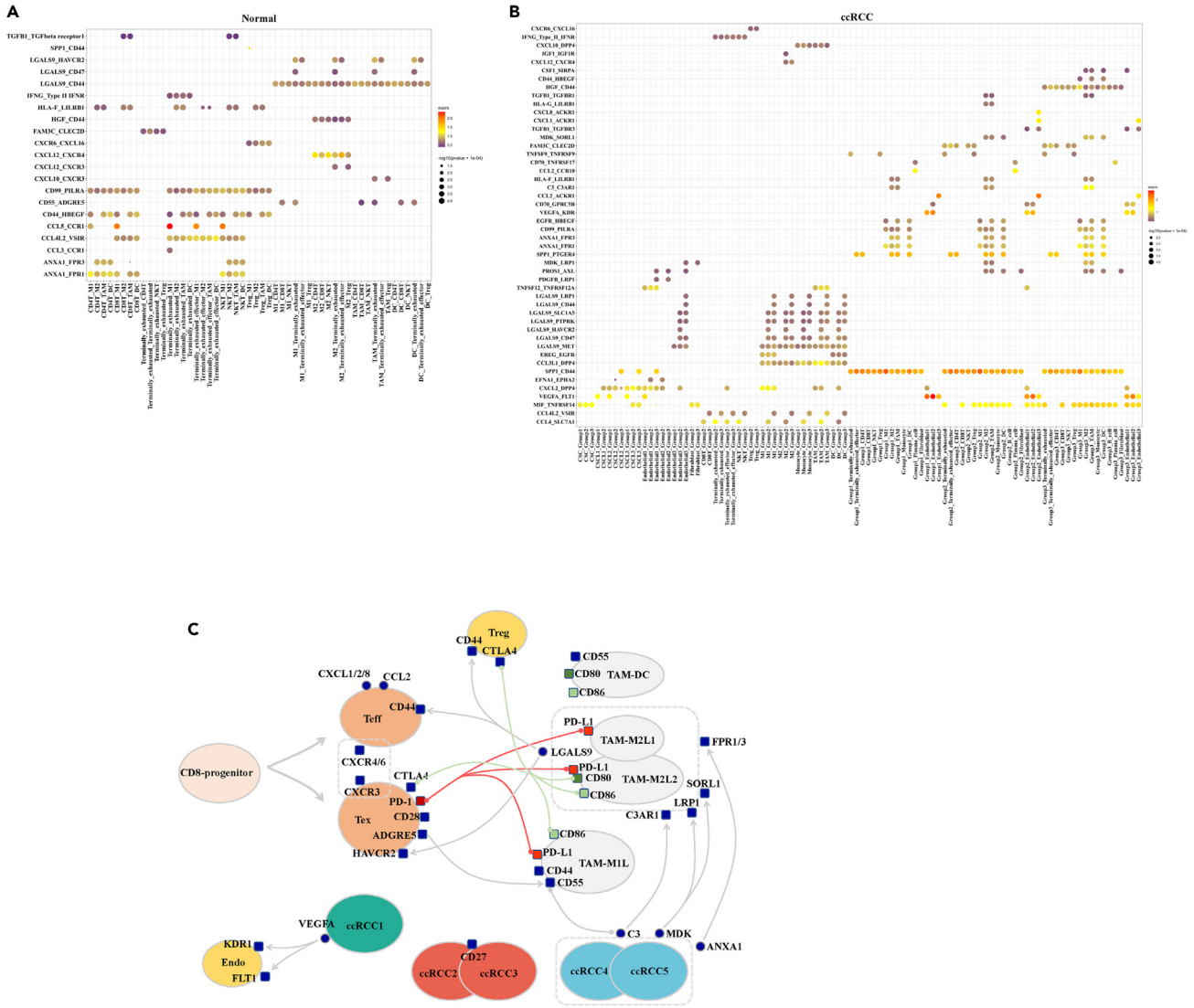


Figure 7. Ligand-receptor interactions in the ccRCC microenvironment

(A) Immune cells interactions. The color represents the expression level of interactions intensity; the size represents the p value of interactions.

(B) Interactions between ccRCC and immune cells. The color represents the expression level of interactions intensity; the size represents the p value of interactions.

(C) Ligand-receptor explanation for immune microenvironment.

cell cluster had distinct TF activation, although they originated from the same CSCLs. FOSL2, KLF2, YY1, and ETS2 are believed to promote tumor angiogenesis,[44–48] which were found enriched in our study, especially in ccRCC1 cluster (group1). TCF12 and SOX4 can possibly increase the risk of tumor proliferation and metastasis,[49–52] which were found enriched in our study, especially in ccRCC4 cluster (group3). These results also show that the pattern of TF activation brought the development of tumor subtypes caused by CSC differentiation.

Currently, systemic drug therapy is the main treatment modality for improving the survival of patients with metastatic ccRCC.[72] The main treatment options of systemic drug therapy include TKI monotherapy and TKI-ICI combination therapy.[72,73] Although systemic drug therapy significantly improves OS and PFS in patients with ccRCC, the response rate to drug therapy is lower than 60% (39% for TKI monotherapy and 59% for TKI-ICI combination therapy).[74] It is worth noting that TKI-ICI combination therapy recommended by EAU RCC guidelines in recent years has become the standard first-line treatment for ccRCC, but targeted monotherapy has gradually relegated to alternative therapy. Indeed, according to the KEYNOTE-426 update study, immune combination therapy is not superior to TKI monotherapy for the IMDC good-risk population. In addition, treatment safety analysis showed that adverse events induced by TKI-ICI combination therapy were more complex, leading to a higher proportion of treatment interruption and discontinuation.[74,75] Therefore, according to different biological characteristics, personalized treatment plans can not only affect the therapeutic efficacy but also avoid the adverse events induced by

excessive treatment. In the present study, group 1 and group 3 clusters were enriched in angiogenesis-related pathways, suggesting that patients in these clusters tend to be sensitive to TKI treatment. As shown recently, patients in angiogenic clusters demonstrated longer PFS with TKI treatment.[9] Meanwhile, group 3 cluster with up-regulated immune-related pathways showed better efficacy for ICI treatment, whereas this was not the case with group 2 cluster. By corollary, ccRCC patients in group 1 were suitable for TKI monotherapy, and ccRCC patients in group 3 were suitable for TKI-ICI combination therapy. Molecular subtyping of tumor differentiation branches can further clarify the nature of different ccRCC tumors, which may be an important basis for solving the mechanism of tumor drug resistance and developing combination therapy, but more in-depth clinical studies are needed.

Tumor-induced T cell dysfunction is a crucial cause of malignant tumor progression and poor immunotherapy efficacy.[76] Previous studies have suggested that T cell exhaustion, identified as a hyporesponsive T cell state, is a key immune escape mechanism for so-called cold tumors.[77] So, are exhausted T cells all at a non-responsive T cell stage? In fact, they are just the terminal state of T cell differentiation. It was found in our study that progenitor-like $CD8^+$ T cells could differentiate into two terminally exhausted states of CD8+ T cells with different clinical outcomes. Increasing evidence has indicated that macrophages are highly plastic in phenotyping and genetics, as macrophages can switch from proinflammatory (M1) to immunosuppressive (M2) phenotype in cancer and other most pathological conditions.[78] M2 macrophages are the major population of leukocytes infiltrating tumors. They secrete multiple growth factors, cytokines, angiogenic factors, and chemokines, some of which inhibit immune-response, causing evasion of tumor immune surveillance.[79] Thus, immunosuppressive TAMs are usually regarded as therapeutic targets. We found that M2Φ cells were positively associated with CD8+Tex cells, which might work through LGALS9/HAVCR2 receptor ligand complex in the cell communication analysis to block the cytotoxicity of CD8+ Tex cells. Previous study also found that the TAMs increased CXCL10 level by binding to its receptor CXCR3 on Tregs and promoted function exhaustion of $CD8^+$ T cells through the elevated expression CTLA4, LAG3, and PDCD1.[66,80] Our findings provide insights into the heterogeneity of the immune microenvironment, and potential therapeutic targets for ccRCC, suggesting that ccRCC patients in group 1 may benefit from TKI monotherapy, whereas TKI-ICI combination therapy may be more applicable in group 3.

In conclusion, we established a ccRCC subtyping system by analyzing cancer cell differentiation trajectories and further validated it in TCGA-KIRC and Checkmate-025 dataset and ccRCC tissue array of our center. We identified three ccRCC molecular subtypes and found that they were associated with different clinical outcomes and distinct treatment-related transcriptomic signatures, suggesting that the classification model may prove to be a useful tool for predicting patient survival and assessing the response of patients to ccRCC therapy. In addition, we described CSC subpopulations and their expression landscapes, believing that they can facilitate the development of potential therapeutic targets against CSCs in ccRCC.

Limitations of the study

This study has some limitations. First, the number (n = 11) of patients with primary ccRCC in our own center is limited, and more tissue samples for scRNA-seq from different institutions are required to further validate and improve the subtyping accuracy. In addition, the ccRCC patients for scRNA-seq were treatment-naive in this study, which might not be fully matched for combination analysis with other data from treated patients. Meanwhile, we had no matched genomics or proteomic data for these samples or datasets, which could not reflect the consistency in terms of both genomic and proteomic features. Further studies can be designed based on our results to investigate the efficacy of promising treatments targeting CSCs on one hand, and on the other hand each ccRCC subpopulation should be investigated in corresponding tumor cells and animals. Finally, well-planned clinical trials are required to validate the efficacy of potential drugs based on this subclassification system.

ETHICS

All procedures obeyed the REMARK and ethical review boards of Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University (XHEC-C-2021-145-1) and the Third Affiliated Hospital of the Second Military Medical University (EHBHKY2020-K-026). Informed consent was obtained from all patients before initiation of the study. All participating patients provided written informed consent.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Preparation of cell suspensions
  - ScRNA sequencing
  - ScRNA-seq data processing
  - Differential gene expression analysis

- Gene set enrichment analysis
- Regulon activity analysis
- Pseudotime trajectory analysis
- Copy number variation (CNV) estimation
- TCGA Kaplan-Meier survival curve analysis
- IHC and cohort survival analysis

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108370.

ACKNOWLEDGMENTS

Funding: This work was sponsored by the National Natural Science Foundation of China (No. 82330094, 82072806, 81974391, 82173265); Leading health talents of Shanghai Municipal Health Commission (2022LJ002); Shanghai Rising-Star Program (23QC1401400); the Natural Science Foundation of Shanghai (23ZR1441300); the Shanghai Municipal Commission of Health and Family Planning (20204Y0042); and Hospital Funded Clinical Research, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (21XHDB06).

AUTHOR CONTRIBUTIONS

Conception and design: Wang Zhou, Xin-gang Cui; Development of methodology: Da Xu, Wen-jin Chen, and Wang Zhou; Acquisition of data: Xiu-wu Pan, Wen-jin Chen, Si-shun Gan, Da Xu, Wen-bin Guan, Lin Li, and Jia-xin Chen; Analysis and interpretation of data: Da Xu, Wen-jin Chen, Wang Zhou, and Xiu-wu Pan; Writing, review, and/or revision of the manuscript: Xiu-wu Pan, Wen-jin Chen, Wang Zhou, and Xin-gang Cui; Administrative, technical, or material support: Si-shun Gan, Wei-jie Chen, Wen-bin Guan, Ke-qin Dong, and Jian-qing Ye; Study supervision: Xin-gang Cui.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

Received: August 3, 2023
Revised: September 3, 2023
Accepted: October 26, 2023
Published: October 30, 2023

REFERENCES

1. Capitanio, U., Bensalah, K., Bex, A., Boorjian, S.A., Bray, F., Coleman, J., Gore, J.L., Sun, M., Wood, C., and Russo, P. (2019). Epidemiology of Renal Cell Carcinoma. Eur. Urol. *75*, 74–84. https://doi.org/10.1016/j.eururo.2018.08.036.
2. Siegel, R.L., Miller, K.D., and Jemal, A. (2020). Cancer statistics, 2020. CA. Cancer J. Clin. *70*, 7–30. https://doi.org/10.3322/caac.21590.
3. Hsieh, J.J., Purdue, M.P., Signoretti, S., Swanton, C., Albiges, L., Schmidinger, M., Heng, D.Y., Larkin, J., and Ficarra, V. (2017). Renal cell carcinoma. Nat. Rev. Dis. Primers *3*, 17009. https://doi.org/10.1038/nrdp.2017.9.
4. Motzer, R.J., Hutson, T.E., Tomczak, P., Michaelson, M.D., Bukowski, R.M., Rixe, O., Oudard, S., Negrier, S., Szczylik, C., Kim, S.T., et al. (2007). Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. N. Engl. J. Med. *356*, 115–124. https://doi.org/10.1056/NEJMoa065044.
5. Motzer, R.J., Escudier, B., McDermott, D.F., George, S., Hammers, H.J., Srinivas, S., Tykodi, S.S., Sosman, J.A., Procopio, G., Plimack, E.R., et al. (2015). Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. N. Engl. J. Med. *373*, 1803–1813. https://doi.org/10.1056/NEJMoa1510665.
6. Motzer, R.J., Hutson, T.E., Cella, D., Reeves, J., Hawkins, R., Guo, J., Nathan, P., Staehler, M., de Souza, P., Merchan, J.R., et al. (2013). Pazopanib versus sunitinib in metastatic renal-cell carcinoma. N. Engl. J. Med. *369*, 722–731. https://doi.org/10.1056/NEJMoa1303989.
7. Qu, L., Wang, Z.L., Chen, Q., Li, Y.M., He, H.W., Hsieh, J.J., Xue, S., Wu, Z.J., Liu, B., Tang, H., et al. (2018). Prognostic Value of a Long Non-coding RNA Signature in Localized Clear Cell Renal Cell Carcinoma. Eur. Urol. *74*, 756–763. https://doi.org/10.1016/j.eururo.2018.07.032.
8. Wei, J.H., Feng, Z.H., Cao, Y., Zhao, H.W., Chen, Z.H., Liao, B., Wang, Q., Han, H., Zhang, J., Xu, Y.Z., et al. (2019). Predictive value of single-nucleotide polymorphism signature for recurrence in localised renal cell carcinoma: a retrospective analysis and multicentre validation study. Lancet Oncol. *20*, 591–600. https://doi.org/10.1016/S1470-2045(18)30932-X.
9. Motzer, R.J., Banchereau, R., Hamidi, H., Powles, T., McDermott, D., Atkins, M.B., Escudier, B., Liu, L.F., Leng, N., Abbas, A.R., et al. (2020). Molecular Subsets in Renal Cancer Determine Outcome to Checkpoint and Angiogenesis Blockade. Cancer Cell *38*, 803–817.e4. https://doi.org/10.1016/j.ccell.2020.10.011.
10. Lindgren, D., Eriksson, P., Krawczyk, K., Nilsson, H., Hansson, J., Veerla, S., Sjölund, J., Höglund, M., Johansson, M.E., and Axelson, H. (2017). Cell-Type-Specific Gene Programs of the Normal Human Nephron Define Kidney Cancer Subtypes. Cell Rep. *20*, 1476–1489. https://doi.org/10.1016/j.celrep.2017.07.043.
11. Zhang, Y., Narayanan, S.P., Mannan, R., Raskind, G., Wang, X., Vats, P., Su, F., Hosseini, N., Cao, X., Kumar-Sinha, C., et al. (2021). Single-cell analyses of renal cell cancers reveal insights into tumor microenvironment, cell of origin, and therapy response. Proc. Natl. Acad. Sci. USA *118*, e2103240118. https://doi.org/10.1073/pnas.2103240118.
12. Ayob, A.Z., and Ramasamy, T.S. (2018). Cancer stem cells as key drivers of tumour progression. J. Biomed. Sci. *25*, 20. https://doi.org/10.1186/s12929-018-0426-4.
13. Tung, M.-C., Lin, Y.W., Lee, W.J., Wen, Y.C., Liu, Y.C., Chen, J.Q., Hsiao, M., Yang, Y.C., and Chien, M.H. (2022). Targeting DRD2 by the antipsychotic drug, penfluridol, retards growth of renal cell carcinoma via inducing stemness inhibition and autophagy-mediated apoptosis. Cell Death Dis. *13*, 400. https://doi.org/10.1038/s41419-022-04828-3.

14. Clark, D.J., Dhanasekaran, S.M., Petralia, F., Pan, J., Song, X., Hu, Y., da Veiga Leprevost, F., Reva, B., Lih, T.S.M., Chang, H.Y., et al. (2020). Integrated Proteogenomic Characterization of Clear Cell Renal Cell Carcinoma. Cell *180*, 207. https://doi.org/10.1016/j.cell.2019.12.026.
15. Motzer, R.J., Robbins, P.B., Powles, T., Albiges, L., Haanen, J.B., Larkin, J., Mu, X.J., Ching, K.A., Uemura, M., Pal, S.K., et al. (2020). Avelumab plus axitinib versus sunitinib in advanced renal cell carcinoma: biomarker analysis of the phase 3 JAVELIN Renal 101 trial. Nat. Med. *26*, 1733–1741. https://doi.org/10.1038/s41591-020-1044-8.
16. McDermott, D.F., Huseni, M.A., Atkins, M.B., Motzer, R.J., Rini, B.I., Escudier, B., Fong, L., Joseph, R.W., Pal, S.K., Reeves, J.A., et al. (2018). Clinical activity and molecular correlates of response to atezolizumab alone or in combination with bevacizumab versus sunitinib in renal cell carcinoma. Nat. Med. *24*, 749–757. https://doi.org/10.1038/s41591-018-0053-3.
17. Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196. https://doi.org/10.1126/science.aad0501.
18. Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., Kang, B., Liu, Z., Jin, L., Xing, R., et al. (2018). Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. Nat. Med. *24*, 978–985. https://doi.org/10.1038/s41591-018-0045-3.
19. Fiedorowicz, M., Khan, M.I., Strzemecki, D., Orzeł, J., Wełniak-Kamińska, M., Sobiborowicz, A., Wieteska, M., Rogulski, Z., Cheda, L., Wargocka-Matuszewska, W., et al. (2020). Renal carcinoma CD105-/CD44- cells display stem-like properties *in vitro* and form aggressive tumors *in vivo*. Sci. Rep. *10*, 5379. https://doi.org/10.1038/s41598-020-62205-6.
20. Bi, K., He, M.X., Bakouny, Z., Kanodia, A., Napolitano, S., Wu, J., Grimaldi, G., Braun, D.A., Cuoco, M.S., Mayorga, A., et al. (2021). Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. Cancer Cell *39*, 649–661.e5. https://doi.org/10.1016/j.ccell.2021.02.015.
21. Gassenmaier, M., Chen, D., Buchner, A., Henkel, L., Schiemann, M., Mack, B., Schendel, D.J., Zimmermann, W., and Pohla, H. (2013). CXC chemokine receptor 4 is essential for maintenance of renal cell carcinoma-initiating cells and predicts metastasis. Stem Cell. *31*, 1467–1476.
22. Ge, Y., Weygant, N., Qu, D., May, R., Berry, W.L., Yao, J., Chandrakesan, P., Zheng, W., Zhao, L., Zhao, K.L., et al. (2018). Alternative splice variants of DCLK1 mark cancer stem cells, promote self-renewal and drug-resistance, and can be targeted to inhibit tumorigenesis in kidney cancer. Int. J. Cancer *143*, 1162–1175.
23. Pode-Shakked, N., Metsuyanim, S., Rom-Gross, E., Mor, Y., Fridman, E., Goldstein, I., Amariglio, N., Rechavi, G., Keshet, G., and Dekel, B. (2009). Developmental tumourigenesis: NCAM as a putative marker for the malignant renal stem/progenitor cell population. J. Cell Mol. Med. *13*, 1792–1808.
24. Merlos-Suárez, A., Barriga, F.M., Jung, P., Iglesias, M., Céspedes, M.V., Rossell, D., Sevillano, M., Hernando-Momblona, X., da Silva-Diz, V., Muñoz, P., et al. (2011). The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. Cell Stem Cell *8*, 511–524. https://doi.org/10.1016/j.stem.2011.02.020.
25. Tostain, J., Li, G., Gentil-Perret, A., and Gigante, M. (2010). Carbonic anhydrase 9 in clear cell renal cell carcinoma: a marker for diagnosis, prognosis and treatment. Eur. J. Cancer *46*, 3141–3148. https://doi.org/10.1016/j.ejca.2010.07.020.
26. Lucarelli, G., Rutigliano, M., Sallustio, F., Ribatti, D., Giglio, A., Lepore Signorile, M., Grossi, V., Sanese, P., Napoli, A., Maiorano, E., et al. (2018). Integrated multi-omics characterization reveals a distinctive metabolic signature and the role of NDUFA4L2 in promoting angiogenesis, chemoresistance, and mitochondrial dysfunction in clear cell renal cell carcinoma. Aging (Albany NY) *10*, 3957–3985. https://doi.org/10.18632/aging.101685.
27. Paret, C., Schön, Z., Szponar, A., and Kovacs, G. (2010). Inflammatory protein serum amyloid A1 marks a subset of conventional renal cell carcinomas with fatal outcome. Eur. Urol. *57*, 859–866.
28. Young, M.D., Mitchell, T.J., Vieira Braga, F.A., Tran, M.G.B., Stewart, B.J., Ferdinand, J.R., Collord, G., Botting, R.A., Popescu, D.M., Loudon, K.W., et al. (2018). Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. Science *361*, 594–599. https://doi.org/10.1126/science.aat1699.
29. Galleggiante, V., Rutigliano, M., Sallustio, F., Ribatti, D., Ditonno, P., Bettocchi, C., Selvaggi, F.P., Lucarelli, G., and Battaglia, M. (2014). CTR2 identifies a population of cancer cells with stem cell-like features in patients with clear cell renal cell carcinoma. J. Urol. *192*, 1831–1841. https://doi.org/10.1016/j.juro.2014.06.070.
30. Pettersen, E.O., Ebbesen, P., Gieling, R.G., Williams, K.J., Dubois, L., Lambin, P., Ward, C., Meehan, J., Kunkler, I.H., Langdon, S.P., et al. (2015). Targeting tumour hypoxia to prevent cancer metastasis. From biology, biosensing and technology to drug development: the METOXIA consortium. J. Enzyme Inhib. Med. Chem. *30*, 689–721.
31. Hamilton, M.J., Young, M., Jang, K., Sauer, S., Neang, V.E., King, A.T., Girke, T., and Martinez, E. (2020). HOTAIRM1 lncRNA is downregulated in clear cell renal cell carcinoma and inhibits the hypoxia pathway. Cancer Lett. *472*, 50–58.
32. Natsuizaka, M., Naganuma, S., Kagawa, S., Ohashi, S., Ahmadi, A., Subramanian, H., Chang, S., Nakagawa, K.J., Ji, X., Liebhaber, S.A., et al. (2012). Hypoxia induces IGFBP3 in esophageal squamous cancer cells through HIF-1α-mediated mRNA transcription and continuous protein synthesis. Faseb. J. *26*, 2620–2630.
33. Claesson-Welsh, L., and Welsh, M. (2013). VEGFA and tumour angiogenesis. J. Intern. Med. *273*, 114–127.
34. D'Andrea, F.P., Safwat, A., Kassem, M., Gautier, L., Overgaard, J., and Horsman, M.R. (2011). Cancer stem cell overexpression of nicotinamide N-methyltransferase enhances cellular radiation resistance. Radiother. Oncol. *99*, 373–378.
35. Ji, Y., Yang, C., Tang, Z., Yang, Y., Tian, Y., Yao, H., Zhu, X., Zhang, Z., Ji, J., and Zheng, X. (2017). Adenylate kinase hCINAP determines self-renewal of colorectal cancer stem cells by facilitating LDHA phosphorylation. Nat. Commun. *8*, 15308–15316.
36. Kang, M.J., Kim, H.P., Lee, K.s., Yoo, Y.D., Kwon, Y.T., Kim, K.M., Kim, T.Y., and Yi, E.C. (2013). Proteomic analysis reveals that CD147/EMMPRIN confers chemoresistance in cancer stem cell-like cells. Proteomics *13*, 1714–1725.
37. Clark, D.J., Dhanasekaran, S.M., Petralia, F., Pan, J., Song, X., Hu, Y., da Veiga Leprevost, F., Reva, B., Lih, T.S.M., Chang, H.Y., et al. (2019). Integrated Proteogenomic Characterization of Clear Cell Renal Cell Carcinoma. Cell *179*, 964–983.e31. https://doi.org/10.1016/j.cell.2019.10.007.
38. Brunelli, M., Eccher, A., Gobbo, S., Ficarra, V., Novara, G., Cossu-Rocca, P., Bonetti, F., Menestrina, F., Cheng, L., Eble, J.N., and Martignoni, G. (2008). Loss of chromosome 9p is an independent prognostic factor in patients with clear cell renal cell carcinoma. Mod. Pathol. *21*, 1–6.
39. Sanjmyatav, J., Junker, K., Matthes, S., Muehr, M., Sava, D., Sternal, M., Wessendorf, S., Kreuz, M., Gajda, M., Wunderlich, H., and Schwaenen, C. (2011). Identification of genomic alterations associated with metastasis and cancer specific survival in clear cell renal cell carcinoma. J. Urol. *186*, 2078–2083. https://doi.org/10.1016/j.juro.2011.06.050.
40. Wettersten, H.I., Aboud, O.A., Lara, P.N., Jr., and Weiss, R.H. (2017). Metabolic reprogramming in clear cell renal cell carcinoma. Nat. Rev. Nephrol. *13*, 410–419. https://doi.org/10.1038/nrneph.2017.59.
41. Xiong, Z., Xiao, W., Bao, L., Xiong, W., Xiao, H., Qu, Y., Yuan, C., Ruan, H., Cao, Q., Wang, K., et al. (2019). Tumor cell "Slimming" regulates tumor progression through PLCL1/UCP1-Mediated lipid browning. Adv. Sci. *6*, 1801862.
42. Qiu, B., Ackerman, D., Sanchez, D.J., Li, B., Ochocki, J.D., Grazioli, A., Bobrovnikova-Marjon, E., Diehl, J.A., Keith, B., and Simon, M.C. (2015). HIF2α-dependent lipid storage promotes endoplasmic reticulum homeostasis in clear-cell renal cell carcinoma. Cancer Discov. *5*, 652–667.
43. Shroff, E.H., Eberlin, L.S., Dang, V.M., Gouw, A.M., Gabay, M., Adam, S.J., Bellovin, D.I., Tran, P.T., Philbrick, W.M., Garcia-Ocana, A., et al. (2015). MYC oncogene overexpression drives renal cell carcinoma in a mouse model through glutamine metabolism. Proc. Natl. Acad. Sci. USA *112*, 6539–6544.
44. Wan, X., Guan, S., Hou, Y., Qin, Y., Zeng, H., Yang, L., Qiao, Y., Liu, S., Li, Q., Jin, T., et al. (2021). FOSL2 promotes VEGF-independent angiogenesis by transcriptionnally activating Wnt5a in breast cancer-associated fibroblasts. Theranostics *11*, 4975–4991.
45. Yu, H., Zheng, J., Liu, X., Xue, Y., Shen, S., Zhao, L., Li, Z., and Liu, Y. (2017). Transcription factor NFAT5 promotes glioblastoma cell-driven angiogenesis via SBF2-AS1/miR-338-3p-mediated EGFL7 expression change. Front. Mol. Neurosci. *10*, 301.
46. Pi, J., Tao, T., Zhuang, T., Sun, H., Chen, X., Liu, J., Cheng, Y., Yu, Z., Zhu, H.H., Gao, W.Q., et al. (2017). A MicroRNA302-367-Erk1/2-Klf2-S1pr1 Pathway Prevents Tumor Growth via Restricting Angiogenesis and Improving Vascular Stability. Circ Res. *120*, 85–98.
47. Yang, D.X., Li, N.E., Ma, Y., Han, Y.C., and Shi, Y. (2010). Expression of Elf-1 and survivin in non-small cell lung cancer and their relationship to intratumoral microvessel density. Chin. J. Cancer *29*, 396–402.
48. Liu, H., Qiu, Y., Pei, X., Chitteti, R., Steiner, R., Zhang, S., and Jin, Z.G. (2020). Endothelial

specific YY1 deletion restricts tumor angiogenesis and tumor growth. Sci. Rep. *10*, 20493.

49. Lee, C.-C., Chen, W.S., Chen, C.C., Chen, L.L., Lin, Y.S., Fan, C.S., and Huang, T.S. (2012). TCF12 protein functions as transcriptional repressor of E-cadherin, and its overexpression is correlated with metastasis of colorectal cancer. J. Biol. Chem. *287*, 2798–2809.
50. Chen, X., Huang, J., and Lü, Y. (2020). High expression of STAT2 in ovarian cancer and its effect on metastasis of ovarian cancer cells. Nan Fang Yi Ke Da Xue Xue Bao *40*, 34–41.
51. Griesmann, H., Mühl, S., Riedel, J., Theuerkorn, K., Sipos, B., Esposito, I., Vanden Heuvel, G.B., and Michl, P. (2021). CUX1 Enhances Pancreatic Cancer Formation by Synergizing with KRAS and Inducing MEK/ERK-Dependent Proliferation. Cancers *13*, 2462.
52. Chen, X., Xu, M., Xu, X., Zeng, K., Liu, X., Pan, B., Li, C., Sun, L., Qin, J., Xu, T., et al. (2020). METTL14-mediated N6-methyladenosine modification of SOX4 mRNA inhibits tumor metastasis in colorectal cancer. Mol. Cancer *19*, 106–116.
53. Fendler, A., Bauer, D., Busch, J., Jung, K., Wulf-Goldenberg, A., Kunz, S., Song, K., Myszczyszyn, A., Elezkurtaj, S., Erguen, B., et al. (2020). Inhibiting WNT and NOTCH in renal cancer stem cells and the implications for human patients. Nat. Commun. *11*, 929.
54. Liu, L., Michowski, W., Inuzuka, H., Shimizu, K., Nihira, N.T., Chick, J.M., Li, N., Geng, Y., Meng, A.Y., Ordureau, A., et al. (2017). G1 cyclins link proliferation, pluripotency and differentiation of embryonic stem cells. Nat. Cell Biol. *19*, 177–188. https://doi.org/10.1038/ncb3474.
55. Zhu, L., Gibson, P., Currle, D.S., Tong, Y., Richardson, R.J., Bayazitov, I.T., Poppleton, H., Zakharenko, S., Ellison, D.W., and Gilbertson, R.J. (2009). Prominin 1 marks intestinal stem cells that are susceptible to neoplastic transformation. Nature *457*, 603–607.
56. Karunanithi, S., Levi, L., DeVecchio, J., Karagkounis, G., Reizes, O., Lathia, J.D., Kalady, M.F., and Noy, N. (2017). RBP4-STRA6 pathway drives cancer stem cell maintenance and mediates high-fat diet-induced colon carcinogenesis. Stem Cell Rep. *9*, 438–450.
57. Horiguchi, H., Endo, M., Kawane, K., Kadomatsu, T., Terada, K., Morinaga, J., Araki, K., Miyata, K., and Oike, Y. (2017). ANGPTL 2 expression in the intestinal stem cell niche controls epithelial regeneration and homeostasis. The EMBO journal *36*, 409–424.
58. Hu, X.-L., Chen, G., Zhang, S., Zheng, J., Wu, J., Bai, Q.R., Wang, Y., Li, J., Wang, H., Feng, H., et al. (2017). Persistent expression of VCAM1 in radial glial cells is required for the embryonic origin of postnatal neural stem cells. Neuron *95*, 309–325.e6.
59. Elizondo, D.M., Andargie, T.E., Lee, C.M., Anderson, W.A., and Lipscomb, M.W. (2017). AIF1 expression regulates differentiation of dendritic cells from hematopoietic stem cell multipotent progenitor subsets and regulates antigen presentation capacity. J. Immunol. *198* (1_Supplement), 202.3. https://doi.org/10.4049/jimmunol.198.Supp.202.3.
60. Zheng, J., Zhang, M., and Weng, H. (2018). Induction of the mitochondrial NDUFA4L2 protein by HIF-1a regulates heart regeneration by promoting the survival of cardiac stem cell. Biochem. Biophys. Res. Commun. *503*, 2226–2233.
61. Hegde, S., Hankey, P., and Paulson, R.F. (2012). Self-renewal of leukemia stem cells in friend virus-induced erythroleukemia requires proviral insertional activation of Spi1 and Hedgehog signaling but not mutation of p53. Stem Cell. *30*, 121–130.
62. Belew, M.S., Bhatia, S., Keyvani Chahi, A., Rentas, S., Draper, J.S., and Hope, K.J. (2018). PLAG1 and USF2 Co-regulate expression of musashi-2 in human hematopoietic stem and progenitor cells. Stem Cell Rep. *10*, 1384–1397.
63. DeLaForest, A., Nagaoka, M., Si-Tayeb, K., Noto, F.K., Konopka, G., Battle, M.A., and Duncan, S.A. (2011). HNF4A is essential for specification of hepatic progenitors from human pluripotent stem cells. Development *138*, 4143–4153.
64. Philip, M., Fairchild, L., Sun, L., Horste, E.L., Camara, S., Shakiba, M., Scott, A.C., Viale, A., Lauer, P., Merghoub, T., et al. (2017). Chromatin states define tumour-specific T cell dysfunction and reprogramming. Nature *545*, 452–456. https://doi.org/10.1038/nature22367.
65. Bieche, I., Vacher, S., Vallerand, D., Richon, S., Hatem, R., De Plater, L., Dahmani, A., Némati, F., Angevin, E., Marangoni, E., et al. (2014). Vasculature analysis of patient derived tumor xenografts using species-specific PCR assays: evidence of tumor endothelial cells and atypical VEGFA-VEGFR1/2 signalings. BMC Cancer *14*, 178. https://doi.org/10.1186/1471-2407-14-178.
66. Zhang, Y., Zuo, C., Liu, L., Hu, Y., Yang, B., Qiu, S., Li, Y., Cao, D., Ju, Z., Ge, J., et al. (2021). Single-cell RNA-sequencing atlas reveals an MDK-dependent immunosuppressive environment in ErbB pathway-mutated gallbladder cancer. J. Hepatol. *75*, 1128–1141. https://doi.org/10.1016/j.jhep.2021.06.023.
67. Moraes, L.A., Ampomah, P.B., and Lim, L.H.K. (2018). Annexin A1 in inflammation and breast cancer: a new axis in the tumor microenvironment. Cell Adh. Migr. *12*, 417–423. https://doi.org/10.1080/19336918.2018.1486143.
68. Yang, H., Li, L., Liu, X., and Zhao, Y. (2021). High Expression of the Component 3a Receptor 1 (C3AR1) Gene in Stomach Adenocarcinomas Infers a Poor Prognosis and High Immune-Infiltration Levels. Med. Sci. Monit. *27*, e927977. https://doi.org/10.12659/MSM.927977.
69. Beksac, A.T., Paulucci, D.J., Blum, K.A., Yadav, S.S., Sfakianos, J.P., and Badani, K.K. (2017). Heterogeneity in renal cell carcinoma. Urol. Oncol. *35*, 507–515. https://doi.org/10.1016/j.urolonc.2017.05.006.
70. Najafi, M., Farhood, B., and Mortezaee, K. (2019). Cancer stem cells (CSCs) in cancer progression and therapy. J. Cell. Physiol. *234*, 8381–8395. https://doi.org/10.1002/jcp.27740.
71. Pan, X.-w., Zhang, H., Xu, D., Chen, J.X., Chen, W.J., Gan, S.S., Qu, F.J., Chu, C.M., Cao, J.W., Fan, Y.H., et al. (2020). Identification of a novel cancer stem cell subpopulation that promotes progression of human fatal renal cell carcinoma by single-cell RNA-seq analysis. Int. J. Biol. Sci. *16*, 3149–3162.
72. Ljungberg, B., Albiges, L., Abu-Ghanem, Y., Bensalah, K., Dabestani, S., Fernández-Pello, S., Giles, R.H., Hofmann, F., Hora, M., Kuczyk, M.A., et al. (2019). European Association of Urology Guidelines on Renal Cell Carcinoma: The 2019 Update. Eur. Urol. *75*, 799–810. https://doi.org/10.1016/j.eururo.2019.02.011.
73. Albiges, L., Powles, T., Staehler, M., Bensalah, K., Giles, R.H., Hora, M., Kuczyk, M.A., Lam, T.B., Ljungberg, B., Marconi, L., et al. (2019). Updated European Association of Urology guidelines on renal cell carcinoma: immune checkpoint inhibition is the new backbone in first-line treatment of metastatic clear-cell renal cell carcinoma. Eur. Urol. *76*, 151–156.
74. Rini, B.I., Plimack, E.R., Stus, V., Gafanov, R., Hawkins, R., Nosov, D., Pouliot, F., Alekseev, B., Soulières, D., Melichar, B., et al. (2019). Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. N. Engl. J. Med. *380*, 1116–1127. https://doi.org/10.1056/NEJMoa1816714.
75. Rini, B.I., Atkins, M.B., Plimack, E.R., Soulières, D., McDermott, R.S., Bedke, J., Tartas, S., Alekseev, B., Melichar, B., Shparyk, Y., et al. (2022). Characterization and Management of Treatment-emergent Hepatic Toxicity in Patients with Advanced Renal Cell Carcinoma Receiving First-line Pembrolizumab plus Axitinib. Results from the KEYNOTE-426 Trial. Eur. Urol. Oncol. *5*, 225–234. https://doi.org/10.1016/j.euo.2021.05.007.
76. Philip, M., and Schietinger, A. (2022). CD8(+) T cell differentiation and dysfunction in cancer. Nat. Rev. Immunol. *22*, 209–223. https://doi.org/10.1038/s41577-021-00574-3.
77. Joyce, J.A., and Fearon, D.T. (2015). T cell exclusion, immune privilege, and the tumor microenvironment. Science *348*, 74–80. https://doi.org/10.1126/science.aaa6204.
78. MacParland, S.A., Liu, J.C., Ma, X.Z., Innes, B.T., Bartczak, A.M., Gage, B.K., Manuel, J., Khuu, N., Echeverri, J., Linares, I., et al. (2018). Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat. Commun. *9*, 4383. https://doi.org/10.1038/s41467-018-06318-7.
79. Noy, R., and Pollard, J.W. (2014). Tumor-associated macrophages: from mechanisms to therapy. Immunity *41*, 49–61. https://doi.org/10.1016/j.immuni.2014.06.010.
80. Sawant, D.V., Yano, H., Chikina, M., Zhang, Q., Liao, M., Liu, C., Callahan, D.J., Sun, Z., Sun, T., Tabib, T., et al. (2019). Adaptive plasticity of IL-10 and IL-35 T cells cooperatively promotes tumor T cell exhaustion. Nat. Immunol. *20*, 724–735. https://doi.org/10.1038/s41590-019-0346-9.
81. Braun, D.A., Hou, Y., Bakouny, Z., Ficial, M., Sant' Angelo, M., Forman, J., Ross-Macdonald, P., Berger, A.C., Jegede, O.A., et al. (2020). Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. Nat. Med. *26*, 909–918.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-BAG3	Abcam	https://www.abcam.cn/products/primary-antibodies/bag3-antibody-ab47124.html
anti-HSPH1	Abcam	https://www.abcam.cn/products/primary-antibodies/hsp105hsp110-antibody-epr4576-ab109624.html
anti-SLC30A1	Invitrogen	https://www.thermofisher.cn/cn/zh/antibody/product/SLC30A1-Antibody-Polyclonal/PA5-37463
anti-TAPP1	Abcam	https://www.abcam.cn/products/primary-antibodies/tapp-1-antibody-ab230181.html
anti-PCNP	Abcam	https://www.abcam.cn/products/primary-antibodies/pcnp-antibody-ab97909.html
anti-BBOX1	novus	https://www.novusbio.com/products/bbox1-antibody_nbp1-88695
anti-SERPINE2	Abcam	https://www.abcam.cn/products/primary-antibodies/serpine2pn-1-antibody-ab154591.html
anti-SEZ6L2	novus	https://www.novusbio.com/products/sez6l2-bsrp-a-antibody_nbp2-38051
anti-MXRA7	Atlas Antibodies	https://www.atlasantibodies.com/products/antigens/control-antigens/prest-antigen/mxra7-antigen-aprest84609/?q=&t=
Biological samples		
Human ccRCC tissues	Xinhua Hospital	N/A
Deposited data		
scRNA-seq raw data	This work, https://bigd.big.ac.cn	https://ngdc.cncb.ac.cn/gsub/submit/bioproject/subPRO020437/
TCGA cohort	GDC Data Portal	https://portal.gdc.cancer.gov/
Checkmate-025 cohort	Braun et al.[81]	https://doi.org/10.1038/s41591-020-0839-y
Software and algorithms		
Cellranger	10X Genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger
R	cran.r-project.org	https://cran.r-project.org/src/base/R-4/R-4.0.2.tar.gz
Seurat	cran.r-project.org	https://cran.r-project.org/web/packages/Seurat/
inferCNV	github.com/broadinstitute/inferCNV	https://github.com/broadinstitute/inferCNV
Monocle3	bioconductor.org	https://cole-trapnell-lab.github.io/monocle3/docs/installation/
pySCENIC	github.com/aertslab/pySCENIC	https://github.com/aertslab/pySCENIC
GSVA	bioconductor.org	https://bioconductor.org/packages/release/bioc/html/GSVA.html
Limma	bioconductor.org	https://bioconductor.org/packages/release/bioc/html/limma.html
ggplot2	cran.r-project.org	https://cran.r-project.org/web/packages/ggplot2/index.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Xin-gang Cui (cuixingang@xinhuamed.com.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The raw sequence data generated in this paper were deposited in the Genome Sequence Archive (GSA) of the National Genomics Data Center, Beijing Institute of Genomics (China National Center for Bioinformation), Chinese Academy of Sciences, and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. This paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Eleven tumor samples were obtained from 11 ccRCC patients who underwent laparoscopic partial nephrectomy (LPN) or laparoscopic radical nephrectomy (LRN) in the Third Affiliated Hospital of the Naval Medical University (Shanghai, China) from October 2018 to September 2019 (The detailed clinical characteristics were presented Table 1). For IHC staining, 286 tumor samples were obtained from ccRCC patients who underwent tumor resection surgery in the Third Affiliated Hospital of Naval Medical University and Xinhua Hospital from May 2012 to November 2017. The clinical features of these patients are summarized in Table S1. This study was approved by the ethics committee of the said university and Xinhua Hospital (EHBHKY2020-K-026; XHEC-C-2021-145-1). All patients in this study provided their informed consent.

ccRCC sequencing data were screened from the Cancer Genome Atlas. The standardized RNA-sequence FPKM and Clinic files were downloaded from the TCGA on March 10, 2021. A total of 527 CCRCC samples with complete clinical follow-up information were obtained.

METHOD DETAILS

Preparation of cell suspensions

The tumor samples surgically removed from the patients were immediately sent to the laboratory within one hour in 4°C liquid medium. All samples were rinsed with phosphate buffer saline (PBS) solution three times and cut into tiny cell clumps of less than 1mm3. Each sample was then transferred into the centrifuge tube with digestion medium with collagenase I (Sigma), collagenase IV (Sigma), DNase IV (Sigma), and hyaluronidase I-S (Sigma) in 0.25% Trypsin (ThermoFisher Scientific) to complete the digestion process in a 37°C thermostatic shaker until the samples were digested to single cells. Next, the incubated samples were filtered with 40-μm nylon meshes (Corning) to remove cell debris. After centrifugation, the cell pellet was resuspended in red blood lysis buffer to eliminate red blood cells. After resuspension with PBS, automatic cytometry (Luna) was used to determine the cell concentration and cell vitality. If the calculated concentration was too high, the liquid volume was adjusted to the appropriate concentration and counted once more. Once the desired cell suspension (1000cells/ul) and cell vitality (>90%) were obtained, they were immediately placed on ice for subsequent GEM preparation and reverse transcription.

ScRNA sequencing

The prepared single-cell suspensions were lysed, and the released RNA was barcoded to scRNA-seq libraries with the Chromium Single Cell 3′ Library and Gel Bead Kit v3 (10x Genomics). The cDNA libraries were sequenced on Illumina Nova6000 with a sequencing depth of at least 100,000 reads per cell and labeled in the human genome (build Grch38) under the operation of CellRanger (10X Genomics). The gene locations were annotated with Ensemble build 95.

ScRNA-seq data processing

Original sequencing data matrices from CellRanger (version 3.0.2) were separately imported to R statistical environment (version 4.0.3), and integrated with Seurat R package (version). To remove the low-expressed genes and low-quality cells, cells that had fewer genes below 200 either over 5000 or with more than 30% mitochondrial reads were depleted, and the genes expressed in at least 3 cells were kept from the original data. Unique molecular identifier (UMI) counts were normalized using the "NormalizeData" function. The 1000 highly variable genes (HVGs) were identified by the "FindVariableGenes" function. Principal component analysis (PCA) was performed on the single-cell expression matrix using the ''RunPCA'' function and top 30 principal components were used for clustering with the Louvain graph-clustering method. Then, we used the "Runharmony" to perform the batch-effect correction for each dataset (https://github.com/immuno-genomics/harmony). Subsequently, a K-nearest-neighbor graph was established based on harmony dimensions for cell clustering analysis. As to sub-clusters (except for cancer cells), the same methods were used for recognizing the variable genes, reducing dimensions and clustering. We use dimensionality reduction and cell clustering method provided by monocle3 downstream to reanalyze tumor cells, and the Batch effect from the samples was eliminated by running the align_cds() function.

Differential gene expression analysis

To identify marker genes of each cluster, we contrasted cells from one cluster to all other cells using the Seurat FindAllMarkers function. For cancer cell populations subdivided by Monocle3, we mapped the grouping information of these cell subgroups back to the Seurat object and calculated the differential genes for the Seurat object that rewrites the grouping information. According to the results of the calculation, the ggplot2 and heatmap packages were used to visually display the heat, violin and bubble maps.

Gene set enrichment analysis

The GSVA R package was applied to estimate pathway activity scores for single cell and visualize it through the heatmap using gene set of C2 and C5 collection obtained from the molecular signature database. The differential activities of pathways were calculated using the limma R package.

Regulon activity analysis

PySCENIC (V1.22) algorithm and Arboreto package GRNBoost2 method were used to build the gene regulatory network (GRN) in all cells combined with cis- Target human motif database (V9). Raw expression data were extracted from the Seurat data. Filtration was performed with default parameters of pySCENIC. Then, grnboost2 method was used to compute GRN. CisTarget databases containing hg38__refseq-r80__10kb_up_and_down_tss.mc9nr.feather and hg38__refseq- r80__10kb_up_and_down_tss.mc9nr.feather and the transcription factor motif annotation database(v9) were used to identify enriched motifs. Aucell function assigns specific values to each cell to represent their regulon activity based on Jensen–Shannon divergence.

Pseudotime trajectory analysis

The Monocle3 algorithm(V0.2.3.0) was used to perform dimensionality reduction, cell clustering and plot trajectories to illustrate the behavioral similarity and transitions among cancer cells. An expression matrix derived from Seurat was used to build a CellDataSet for Monocle pipeline, and partition cells into supergroups after dimensionality reduction.

Copy number variation (CNV) estimation

To infer CNV patterns from the scRNA-seq data, we used the inferCNV (V1.6.0) method with recommended parameters for 10X data to illustrate the diverse patterns of chromosome CNV in tumor cell clusters. The Fibroblast, endothelial and normal epithelial cells were used as the reference with low CNV.

TCGA Kaplan-Meier survival curve analysis

The downloaded TCGA CCRCC data were normalized and then integrated. For the integrated data set, Kaplan-Meier survival curves of different subtype gene sets in the data set were drawn using the survival package.

IHC and cohort survival analysis

The IHC was performed according to protocol. Antibody information were as following: BAG3 (Abcam, ab47124), HSPH1 (Abcam, ab109624), SLC30A1 (Invitrogen, PA5-37463), TAPP1 (Abcam, ab230181), PCNP (Abcam, ab97909), BBOX1 (novus, NBP1-88695), SERPINE2 (Abcam, ab154591), SEZ6L2 (novus, NBP2-38051) and MXRA7 (Atlas Antibodies, HPA044819). IHC score and Kaplan-Meier survival curve were used to evaluate the relationship between the expression of genes and prognosis. The survival status and duration of the patients were used to draw Kaplan-Meier curve ($P < 0.05$ indicating a statistically significant difference).

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were performed using R software (version 4.0.2). In Figures S8B and S11B, two-sided Wilcoxon test was used to compare signature scores or ESTIMATE-inferred immune and tumor scores between different cell groups and bulk RNA-Seq sample groups. In Figures 4E, 4I, and 4J, Log-rank test was used to evaluate survival differences. Detailed descriptions of the statistical tests performed for individual analysis are provided in the Figure legends and STAR Methods.