# SUPPLEMENTARY FIGURES



**Supplementary Figure 1. Quality control and sample integration of scRNA-seq and scATAC-seq data.** (A) Violin plots showing quality control indices of scRNA-seq data, including the number of genes detected per cell, count of reads per cell, percentage of mitochondrial genes detected per cell, percentage of ribosomal genes detected per cell, and percentage of hemoglobin genes detected per cell. (B) Sample integration of scRNA-seq data. (C) Violin plots showing quality control indices of scATAC-seq data, including nucleosome signal scores per cell, TSS enrichment scores per cell, blacklist ratio, and fractions of reads in peaks. (D) TSS enrichment score (left) and fragment length (right) of each sample in scATAC-seq data. (E) Sample integration of scATAC-seq data.



SPI1

S100A4

CTSS

STX11

LYZ

CYBB

AC015987.1 ADGRE1

VCAN

TMSB4X

**RNF122** 

CLEC4A

AC135782.1

PIK3AP1

CLEC4D

CLEC7A

RGS18

FCER1G

MS4A7

TNFRSF1B

## В

# ccRCC1: 325 cells; 3 programs



## EIF1 PTGES3L BDKRB1 AC010336.1 ZNF622 AP002793.1 AC010542.2 C6orf47.AS1 RPS10 MYLPF FEN1 CHODL **ZNF792** ZNF823 ZNF554 SCNN1G AC024896.1 AC009812.4 ANXA3 TM4SF19 ADPRM AC072054.1 CASC1 AC097359.2 TAGLN ABTB2 APOBEC3H ZNF790 PGBD1 LINC01063 ZNF407 CSAD AC002558.3 POLR2J3 CMYA5 ANKRD FAM153CP MALAT1 ATR NEAT1 LUC7L3 LINC01320 FTX TARBP1 ZKSCAN1 AL109615.2 HMGN3.AS1 AKAP9 AC048341.1 TTC14 UGT1A1 AC006994.2 TENM AL391117.1 COL27A1 LINC01134 PAX8 AC245060.5 AC073195.1 ANKRD36C PECAM1 IFI30 FGL2 ARHGAP30 ARRB2 SAT1 TNFSF13B CD48 PRKCB RASGRP2 AC090559.1 ZSWIM6

IRAK3

ITGB2

KYNU

ccRCC1\_cNMF\_1; ccRCC1\_cNMF\_2; ccRCC1\_cNMF\_3

# ccRCC2: 332 cells; 3 programs

	ccR0	CC2_cNI	MF_1; c	cRCC2_	cNMF_	2; ccRCC2_cNMF_3
ARL4C	BPTF	PRRC2C	IDS	TNFAIP3	NTM	
STK17B	AHNAK	MALAT1	TGFBI	PLOD2	ZFP36L2	
PTPN22	SPOCK2	RCAN3	CD2	SLA	CENPC	
EPC2	BTG2	RAB11FIP1	TFPI	APOBEC3H	BGN	
ZEB2	BTG1	ZFP36L1	YWHAZ	TNIK	BCL11B	
NAT8	TMEM176A	TMEM176B	MIF	CYP27A1	ANXA4	
CUBN	TMEM37	FXYD2	TNFAIP6	CYB5A	BBOX1	
CXCL14	ACAA2	CLU	UGT2B7	SLC17A3	PEBP1	
ACMSD	SMIM24	SLC3A1	CDH16	LY6E	GAL3ST1	
IMPA2	RTN4	ACE2	MSRB1	AMN	AOC1	
BCL2A1	PLAUR	IFI30	AIF1	HLA.DRA	LILRB2	
FCER1G	CD83	CD33	LST1	PLEK	HLA.DPB1	
SRGN	SMIM25	TYROBP	CD74	POU2F2	CD86	
0	0		- 214			

VSIR

HLA.DQA1

CBAR

FGL2

## ccRCC3: 699 cells; 4 programs ccRCC3\_cNMF\_1; ccRCC3\_cNMF\_2; ccRCC3\_cNMF\_3; ccRCC3\_cNMF\_4 1 All TMEM176B TMEM176A KRT18 AKR1C3 CLU S100A1 CYB5A CXCL14 NNMT IMPA2 ALDH1A1 SPP1 Vold B ANXA KRT8 PDZK1IP1 EXYD2 PDZK1 SNHG2 SOD2 NAPSA ACSM2B CLDN4 SPINT2 ė MGST ENO1 CYSTM CD24 RACK TMEM37 TPI1 CST7 TMSB4X ARHGDIR CCI 5 COROLA NKG7 CD7 LCK CD2 GZMA CD3D S100A4 CTSV diam'r. CD52 TRBC2 LAG3 TRAC CD3G HCST CD247 SRGN SPN TRBC GATA3 GMFG CRTA RUNX CD3 CD53 COL1A2 FN1 COL1A' MEDAG IGFBP2 SPARC FAP CPE IGFBP7 BGN VEGFC MFAP4 CTHRC1 CLEC3B COL3A1 SFRP2 TIMP1 CREB3L1 COL 5A1 GGT5 PLAU AC110285.3 194 COL10A1 AC020910 MC3.AS1 RAB39A KRTAP5.10 PPP1R26.AS SLC28A3 HAS2.AS1 LILRB2 CYBB AIF1 CD86 LILRB4 CSTA AC004847 020656 LYZ LY86 na isan pipina nganan na isang pina ng IF130 CTSS FCN1 MS4A6A AL353804.1 OLIG1 KCNE1B VSTM1 SIGLEC9 CLEC4G AP000919.4 EME1 P2RX1 AC134312.6 AC069368.2 LINC00677 AL132709.8 KRTAP5.8 AC116535.1 AL592071.

HLA.DPA1

MS4A7

CCL3L1

Supplementary Figure 2. Identification of functionally heterogeneous cancer cells in ccRCC based on cNMF algorithm. (A) Stability and error curve for inferring gene expression program numbers. The blue circle indicates the optimal number of gene expression program per sample, while the heatmap shows Euclidean distance of programs across replicates. (B) Heatmap showing characteristic gene expression of programs in each sample. Right box shows the characteristic gene of different programs.

0

-1



Supplementary Figure 3. Molecular characteristics of functionally heterogeneous cancer cells of different meta programs in ccRCC. (A) Transcriptional expression levels and chromatin accessibility of CA9 and NDUFA4L2 in different meta programs of cancer cells. (B) Heatmap showing large-scale copy number variations for cancer cells. Red indicates amplification while blue indicates deletion. T cells were used as reference cells to infer copy number variations of cancer cells.



Supplementary Figure 4. Expression profiles of signature genes in heterogeneous cancer cells of three meta programs and in normal kidney cells. (A) Chromatin accessibility of REGA1, IL10RA and COL1A1 in three meta programs. (B–D) Spatial expression patterns for REGA1, IL10RA and COL1A1 in stRNA-seq data. (E) UMAP embedding of cells from normal kidneys in scRNA-seq data. (F) Dot plot for expression patterns of marker genes for different cell types of normal kidneys in scRNA-seq data. (G) Expression levels of REGA1, IL10RA and COL1A1 for different cell types from normal kidneys in scRNA-seq data.



Supplementary Figure 5. Characteristics of heterogeneous cancer cells and associated transcription factor regulatory **networks.** (A) Estimated proportions and differential analysis of relative abundance of functionally heterogeneous cancer cells calculated by CIBERSORTx. (B) Correlations between infiltration level of heterogeneous cancer cells of disparate meta programs calculated by CIBERSORTx and ssGSEA scores calculated by characteristic genes of meta programs. (C) Functional enrichment analysis results of downstream target genes of specific transcription factors in disparate meta programs.



Supplementary Figure 6. Expression level of therapeutic target genes in normal kidneys based on scRNA-seq data.