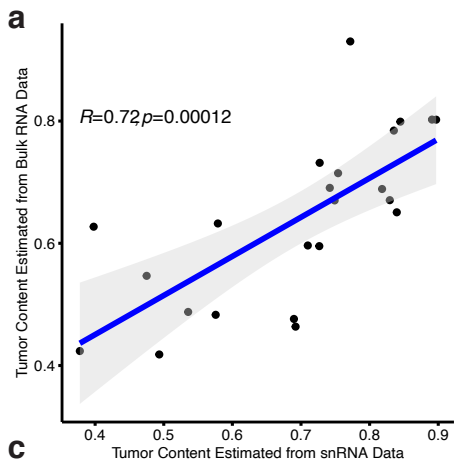
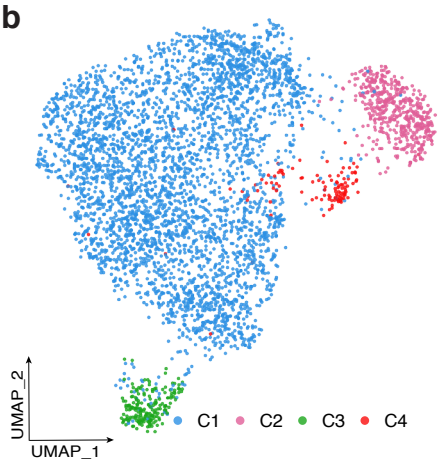
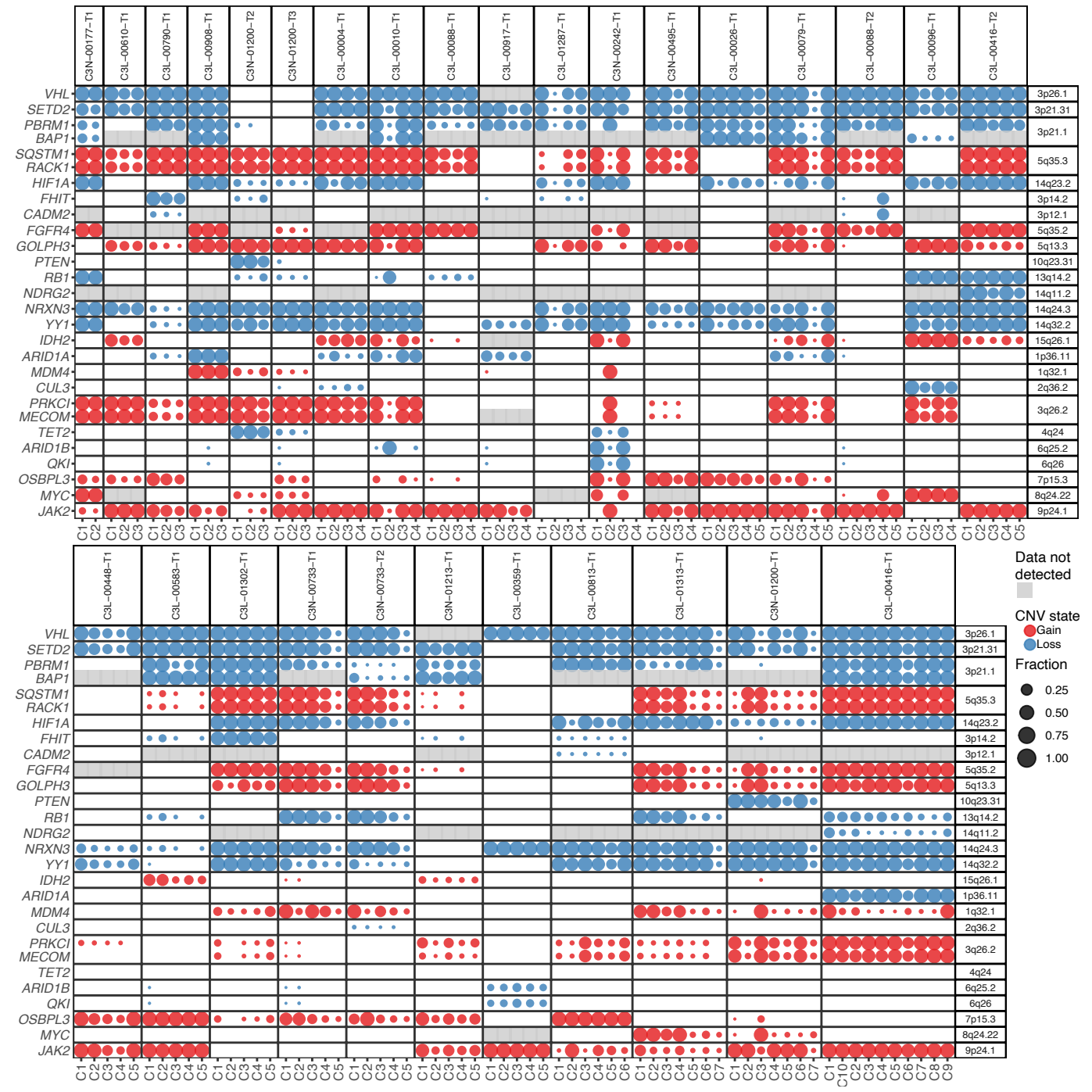


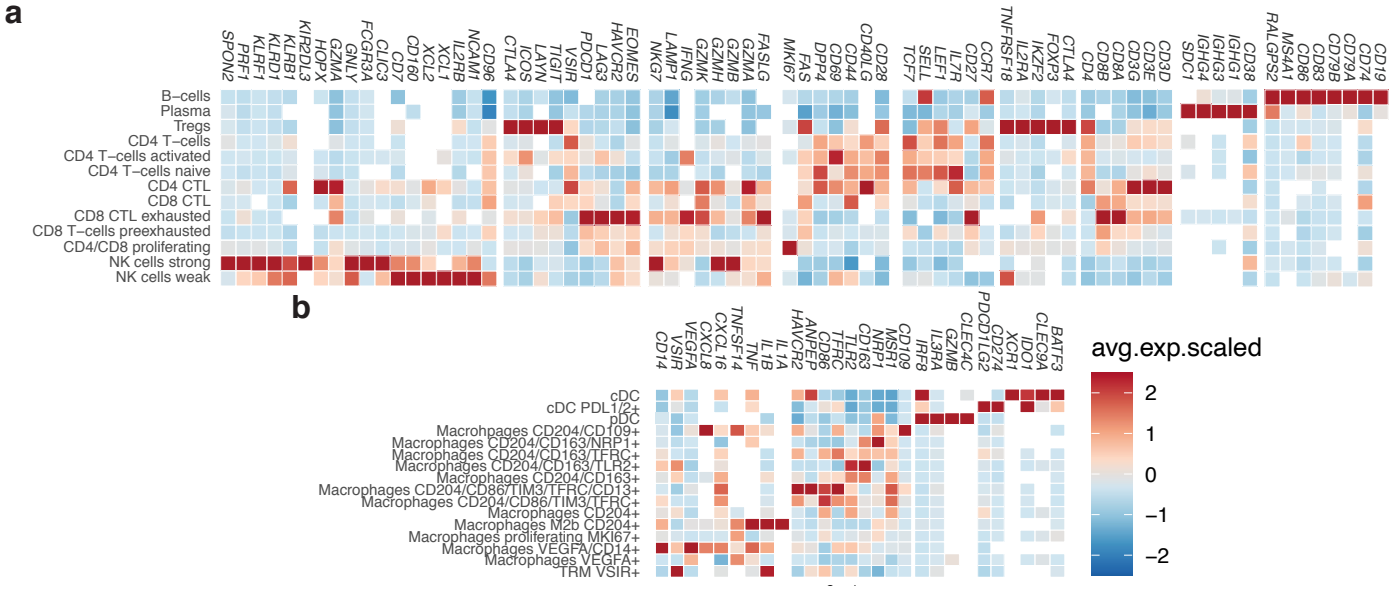


d

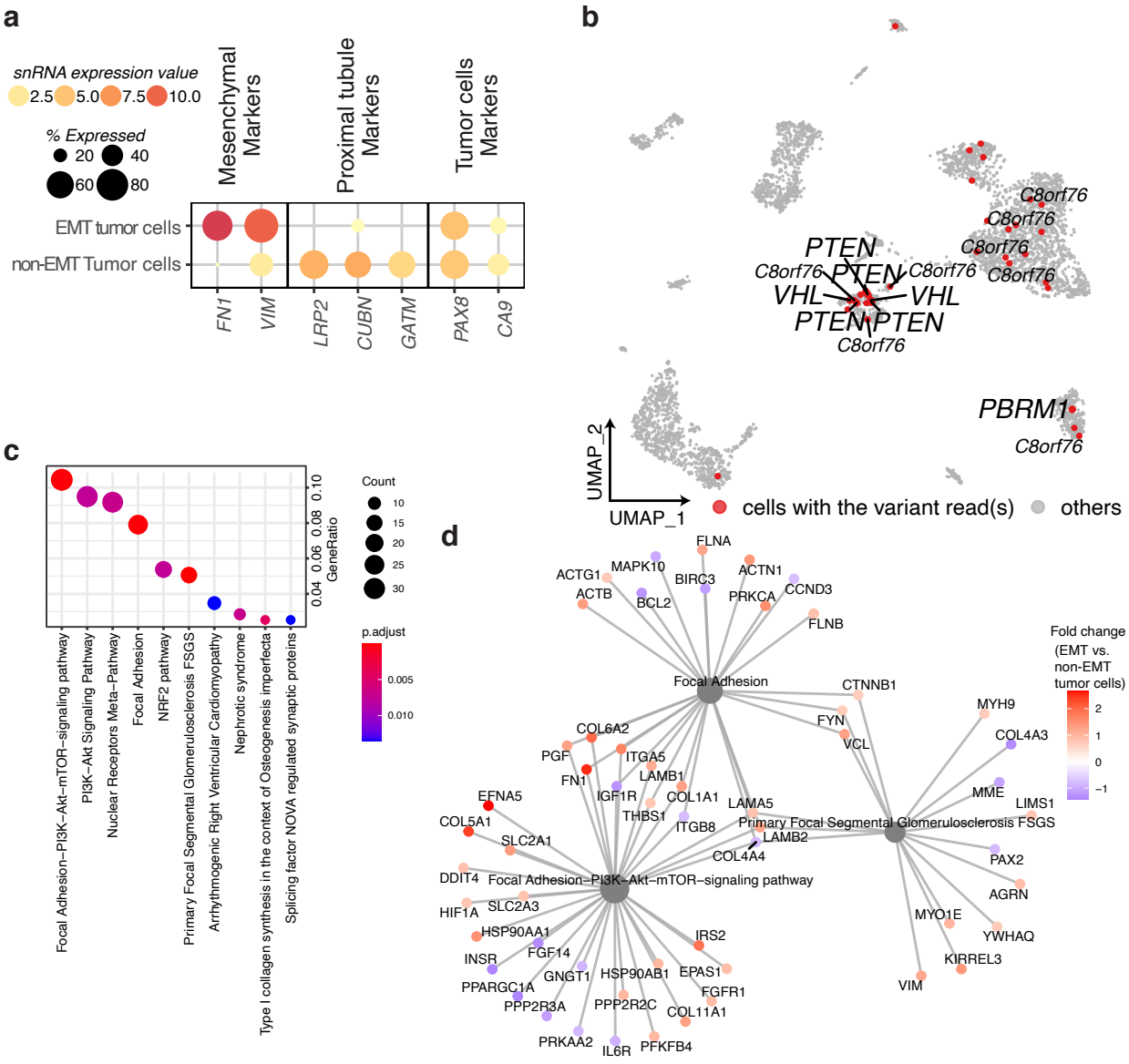
No. Samples		5q gain		
		Complete	Inter-mingled	Minimal
3p deletion	Complete	9	1	6
	Intermingled	3	7	1
	Minimal	2	0	0



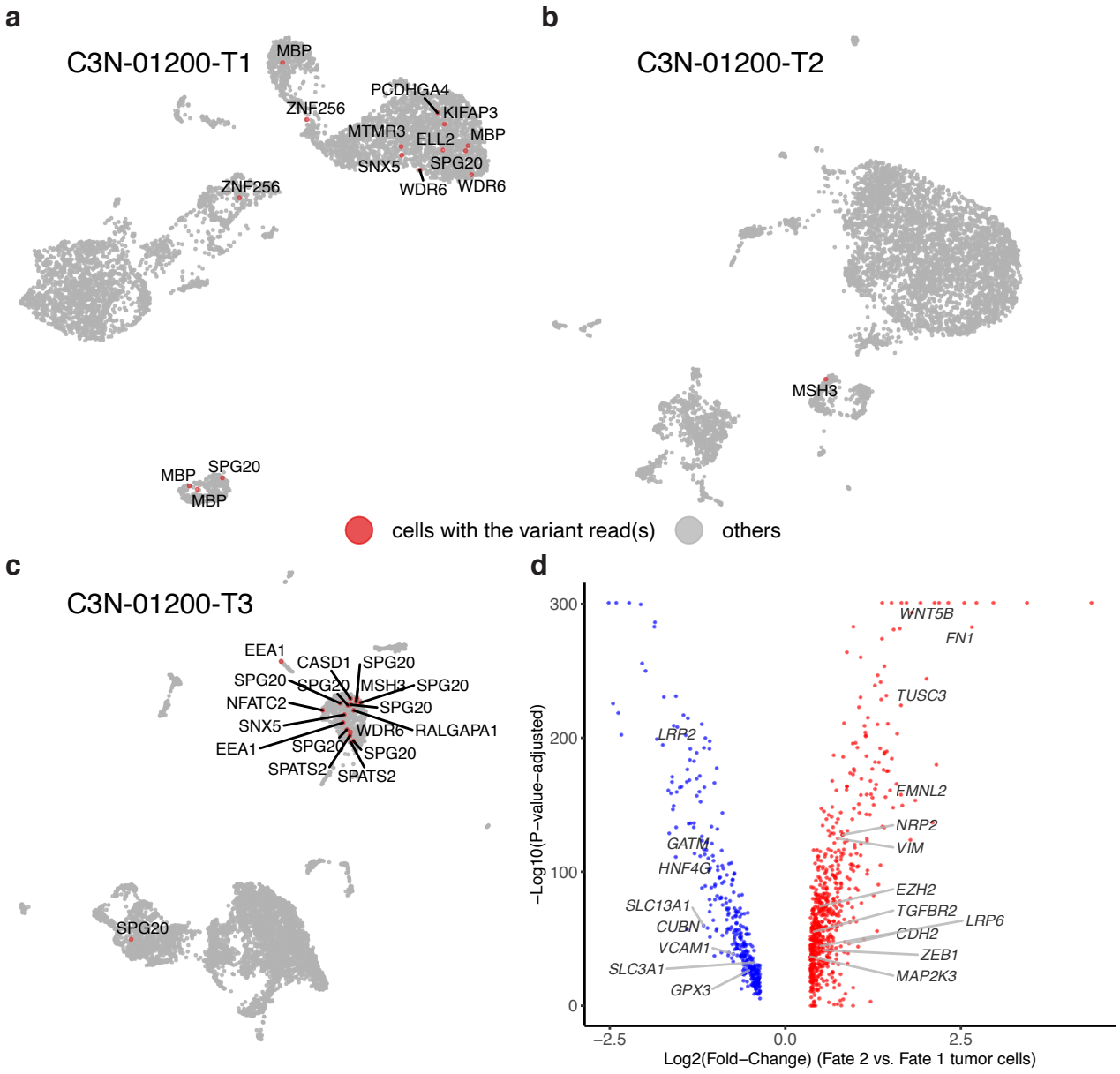
Extended Data Figure 1. Related to Figure 1.
a, Scatter plot showing the tumor cell percentages estimated by the snRNA data (x-axis) were significantly correlated with the tumor purity estimates by ESTIMATE using bulk RNA data (y-axis) among 23 ccRCC tumors. b, UMAP visualization of tumor cells from tumor C3L-00010-T1 (4,377 cells), colored by the tumor subcluster id. c, Bubble plot showing the CNV status and fraction of tumor subclusters (n = 134). d, Table showing the number of tumor samples with different combinations of 3p and 5q CNV status.


a
SPON2
PRF1
KLRF1
KLRD1
KLRB1
KIR2DL3
HOPX
GZMA
GNLY
FCGR3A
CLIC3
CD7
CD160
XCL2
XCL1
IL2RB
NCAM1
CD96
CTLA4
ICOS
LAYN
TIGIT
VSIR
PDCD1
LAG3
HAVCR2
EOMES
NKG7
LAMP1
IFNG
GZMK
GZMH
GZMB
GZMA
FASLG
MKI67
FAS
DPP4
CD69
CD44
CD40LG
CD28
TCF7
SELL
LEF1
IL7R
CD27
CCR7
TNFRSF18
IL2RA
IKZF2
FOXP3
CTLA4
CD4
CD8B
CD8A
CD3G
CD3E
CD3D
SDC1
IGHG4
IGHG3
IGHG1
CD38
RALGPS2
MS4A1
CD86
CD83
CD79B
CD79A
CD74
CD19
B–cells
Plasma
Tregs
CD4 T–cells
CD4 T–cells activated
CD4 T–cells naive
CD4 CTL
CD8 CTL
CD8 CTL exhausted
CD8 T–cells preexhausted
CD4/CD8 proliferating
NK cells strong
NK cells weak
b
CD14
VSIR
VEGFA
CXCL8
CXCL16
TNFSF14
TNF
IL1B
IL1A
HAVCR2
ANPEP
CD86
TFRC
TLR2
CD163
MSR1
CD109
IRF8
IL3RA
GZMB
CLEC4C
CD274
PDCD1LG2
XCR1
IDO1
CLEC9A
BATF3
cDC
cDC PDL1/2+
pDC
Macrohpages CD204/CD109+
Macrophages CD204/CD163/NRP1+
Macrophages CD204/CD163/TFRC+
Macrophages CD204/CD163/TLR2+
Macrophages CD204/CD163+
Macrophages CD204/CD86/TIM3/TFRC/CD13+
Macrophages CD204/CD86/TIM3/TFRC+
Macrophages CD204+
Macrophages M2b CD204+
Macrophages proliferating MKI67+
Macrophages VEGFA/CD14+
Macrophages VEGFA+
TRM VSIR+
avg.exp.scaled
2
1
0
−1
−2

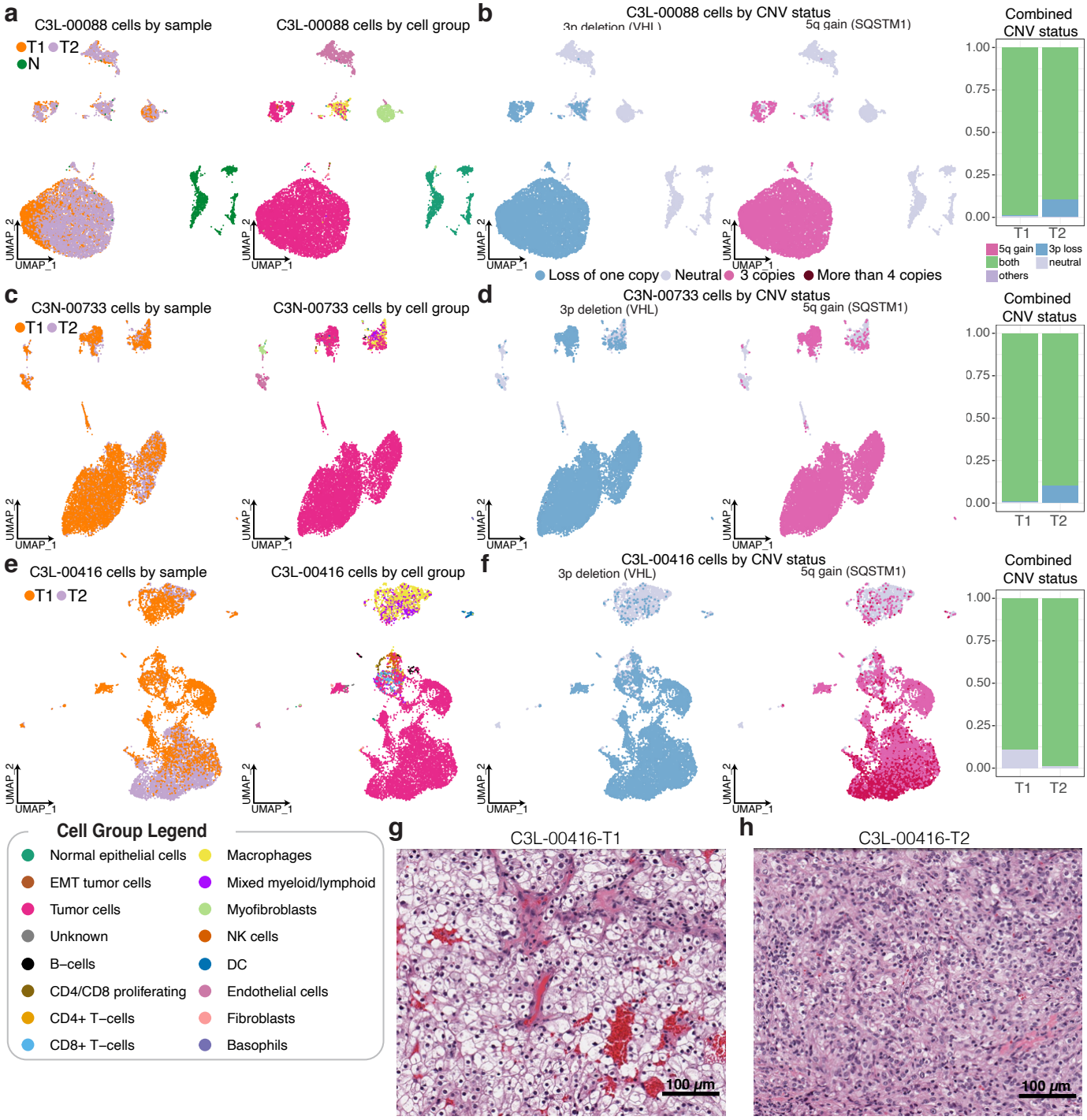
Extended Data Figure 2. Marker gene expression for assigning immune cell states.
a, Heatmap showing the snRNA expression of known marker genes (x-axis) in various lymphoid cell types (y-axis). Expression was averaged across cells within each cell type and then scaled across all lymphoid cell types. Expression is shown only for cell types expressing a given gene in at least 1% of cells. b, Expression of known marker genes (x-axis) in various myeloid cells (y-axis). Expression was averaged across cells within each cell type and then scaled across all myeloid cell types. Expression is shown only for cell types expressing a given gene in at least 1% of cells.


a
snRNA expression value
2.5 5.0 7.5 10.0
% Expressed
20 40 60 80
Mesenchymal Markers
Proximal tubule Markers
Tumor cells Markers
EMT tumor cells
non-EMT Tumor cells
FN1
VIM
LRP2
CUBN
GATM
PAX8
CA9
b
C8orf76
PTEN
VHL
PBRM1
UMAP_2
UMAP_1
cells with the variant read(s)
others
c
Count
10 15 20 25 30
p.adjust
0.005
0.010
GeneRatio
0.04 0.06 0.08 0.10
Focal Adhesion−PI3K−Akt−mTOR−signaling pathway
PI3K−Akt Signaling Pathway
Nuclear Receptors Meta−Pathway
Focal Adhesion
NRF2 pathway
Primary Focal Segmental Glomerulosclerosis FSGS
Arrhythmogenic Right Ventricular Cardiomyopathy
Nephrotic syndrome
Splicing factor NOVA regulated synaptic proteins
Type I collagen synthesis in the context of Osteogenesis imperfecta
d
Fold change (EMT vs. non-EMT tumor cells)
2 1 0 −1
Focal Adhesion
Primary Focal Segmental Glomerulosclerosis FSGS
Focal Adhesion−PI3K−Akt−mTOR−signaling pathway
FLNA
ACTG1
MAPK10
BIRC3
ACTN1
ACTB
BCL2
PRKCA
CCND3
FLNB
CTNNB1
FYN
VCL
MYH9
COL4A3
COL6A2
PGF
ITGA5
FN1
LAMB1
IGF1R
COL1A1
LAMA5
MME
LIMS1
EFNA5
THBS1
ITGB8
LAMB2
COL5A1
SLC2A1
COL4A4
PAX2
DDIT4
AGRN
HIF1A
SLC2A3
MYO1E
HSP90AA1
YWHAQ
IRS2
FGF14
KIRREL3
INSR
GNGT1
HSP90AB1
EPAS1
VIM
PPARGC1A
FGFR1
PPP2R3A
PPP2R2C
COL11A1
PRKAA2
IL6R
PFKFB4

Extended Data Figure 3. Related to Figure 2.
a, Dot plot showing the snRNA expression of mesenchymal marker genes and proximal tubule marker genes in EMT tumor cells and non-EMT tumor cells in tumor C3L-00079-T1. b, UMAP visualization of cells in tumor C3L-00079-T1 (4,350 cells). Red dots denote cells with reads mapped with somatic mutation(s) (*FTH1* mutation not shown to avoid crowdedness). Selected mutations were labeled with the gene symbol, including mutations in *VHL*, *PTEN*, *PBRM1*, and *C8orf76* (same mutation for the same gene across cells). c, Dot plot showing the top ten pathways enriched in the genes differentially expressed between the EMT tumor cells and the non-EMT tumor cells in tumor C3L-00079-T1 using over-representation analysis. d, Network plot showing genes differentially expressed between the EMT tumor cells and non-EMT tumor cells within the top three most significantly enriched pathways (shown in b) in tumor C3L-00079-T1.

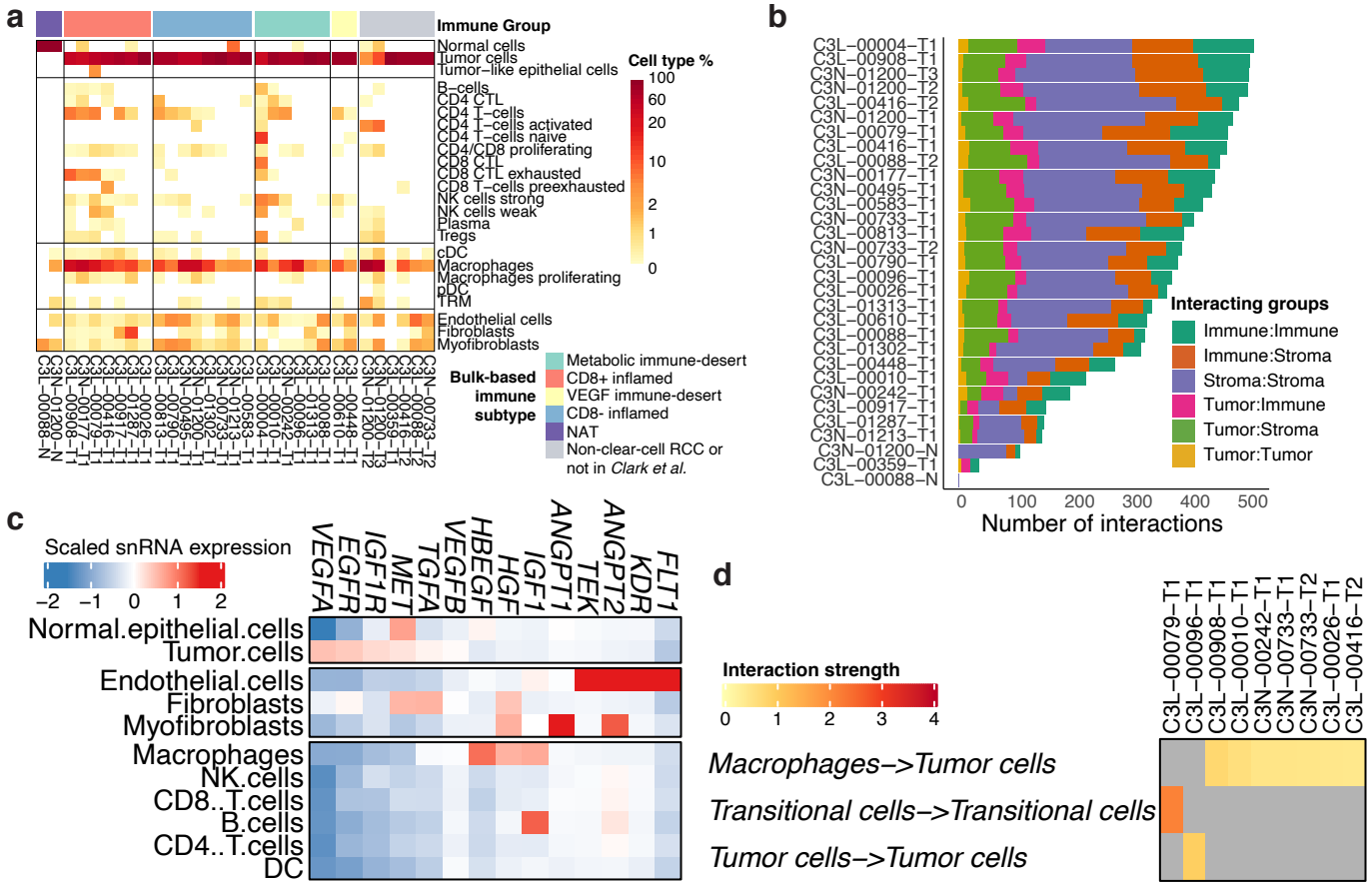

a
C3N-01200-T1
MBP
PCDHGA4
ZNF256
KIFAP3
MTMR3
ELL2
MBP
SNX5
SPG20
ZNF256
WDR6
WDR6
MBP
SPG20
MBP
b
C3N-01200-T2
MSH3
cells with the variant read(s)
others
c
C3N-01200-T3
EEA1
CASD1
SPG20
SPG20
SPG20
MSH3
SPG20
NFATC2
SPG20
SNX5
WDR6
RALGAPA1
EEA1
SPG20
SPG20
SPATS2
SPATS2
SPG20
d
300
200
100
0
−Log10(P−value−adjusted)
WNT5B
FN1
TUSC3
LRP2
FMNL2
NRP2
VIM
GATM
HNF4G
EZH2
SLC13A1
TGFBR2
CUBN
LRP6
CDH2
VCAM1
ZEB1
SLC3A1
MAP2K3
GPX3
−2.5
0.0
2.5
Log2(Fold−Change) (Fate 2 vs. Fate 1 tumor cells)

Extended Data Figure 4. snRNA analysis of the tumors from patient C3N-01200.
a-c, UMAP visualization of cells from tumor segment a, T1 (5,090 cells), b, T2 (5,134 cells), and c, T3 (4,974 cells) of patient C3N-01200. Red dots denote cells with reads mapped with somatic mutation(s). d, Scatter plot showing differentially expressed between Fate 2 tumor cells (EMT tumor cells) vs. Fate 1 tumor cells from patient C3N-01200. Texts on the right denote known mesenchymal genes. Texts on the left denote known marker genes for proximal tubule cells.

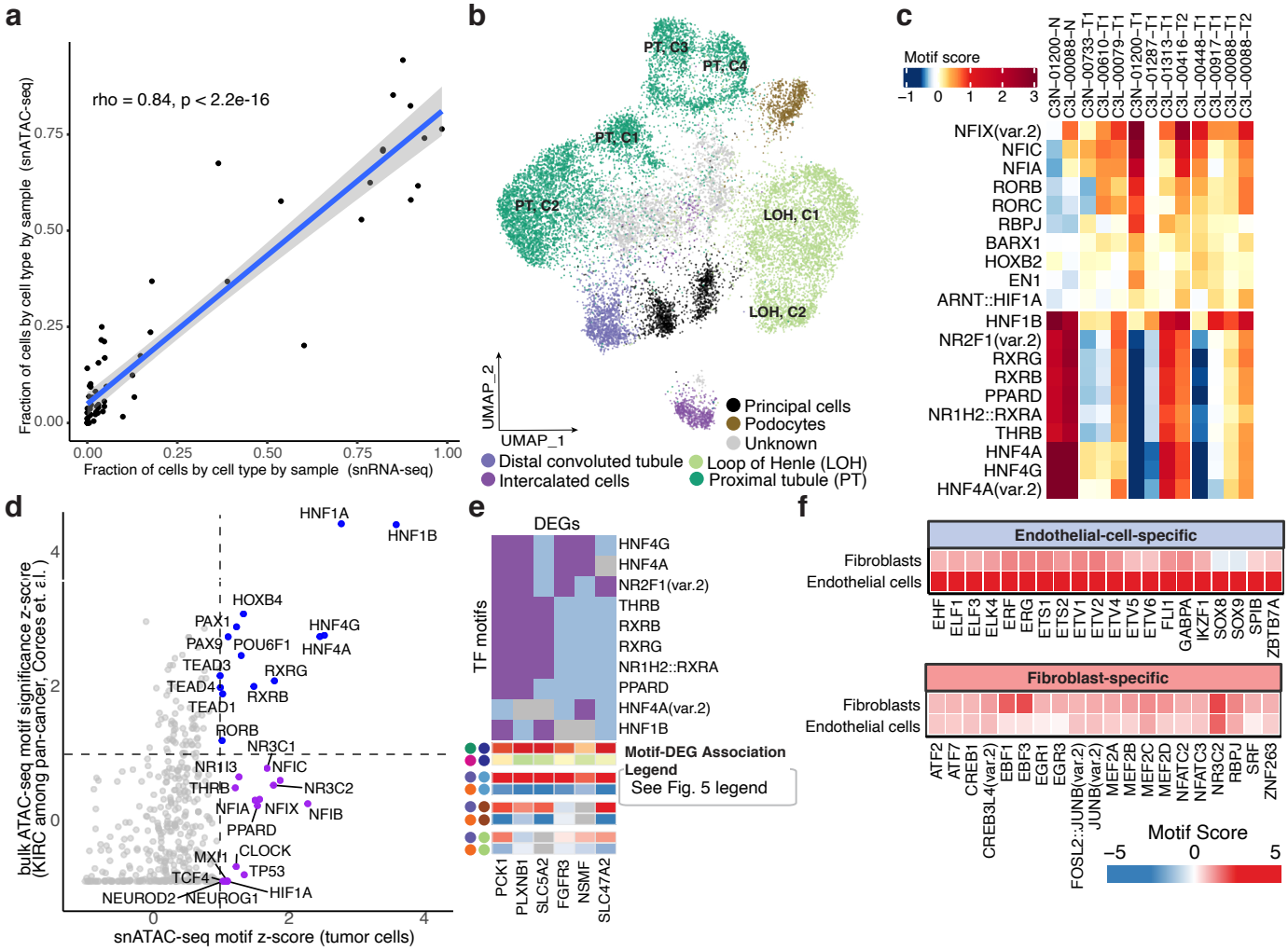

a
C3L-00088 cells by sample
T1 T2
N
C3L-00088 cells by cell group
b
C3L-00088 cells by CNV status
3p deletion (VHL)
5q gain (SQSTM1)
Combined CNV status
1.00
0.75
0.50
0.25
0.00
T1
T2
5q gain
3p loss
both
neutral
others
Loss of one copy
Neutral
3 copies
More than 4 copies
UMAP_1
UMAP_2
c
C3N-00733 cells by sample
T1 T2
C3N-00733 cells by cell group
d
C3N-00733 cells by CNV status
3p deletion (VHL)
5q gain (SQSTM1)
Combined CNV status
e
C3L-00416 cells by sample
T1 T2
C3L-00416 cells by cell group
f
C3L-00416 cells by CNV status
3p deletion (VHL)
5q gain (SQSTM1)
Combined CNV status
Cell Group Legend
Normal epithelial cells
EMT tumor cells
Tumor cells
Unknown
B–cells
CD4/CD8 proliferating
CD4+ T–cells
CD8+ T–cells
Macrophages
Mixed myeloid/lymphoid
Myofibroblasts
NK cells
DC
Endothelial cells
Fibroblasts
Basophils
g
C3L-00416-T1
100 μm
h
C3L-00416-T2
100 μm

Extended Data Figure 5. snRNA analysis of the multiple-segment tumors from case C3L-00088, C3N-00733, and C3L-00416.

a, UMAP visualization of cells (n = 11,202) from patient C3L-00088 T1, T2 tumor segments, and NAT (green), colored by sample (left) and by cell group (right, color key at bottom). b, UMAP visualization of cells from patient C3L-00088, colored by CNV status of VHL (representing 3p) and SQSTM1 (representing 5q) genes; Bar plot showing the fractions of tumor cells in tumor C3L-00088 T1 and T3, colored by the combined CNV status for 3p and 5q gain. c, UMAP visualization of cells (n = 9,138) from patient C3N-00733 T1 and T2 tumor segments, colored by sample (left) and by cell group (right). d, UMAP visualization of cells from patient C3N-00733, colored by CNV status of *VHL* and *SQSTM1* genes; Bar plot showing the fractions of tumor cells in tumor C3L-00088 T1 and T3, colored by the combined CNV status for 3p and 5q gain. e, UMAP visualization of cells (n = 9,842) from patient C3L-00416 T1 and T2 tumor segments, colored by sample (left) and by cell group (right). f, UMAP visualization of cells from patient C3L-00416, colored by CNV status of *VHL* and *SQSTM1* genes; Bar plot showing the fractions of tumor cells in tumor C3L-00088 T1 and T3, colored by the combined CNV status for 3p and 5q gain. g-h, H&E image of a cross-section of tumor g, C3L-00416-T1 and h, C3L-00416-T2. Scale bar, 100 μm.

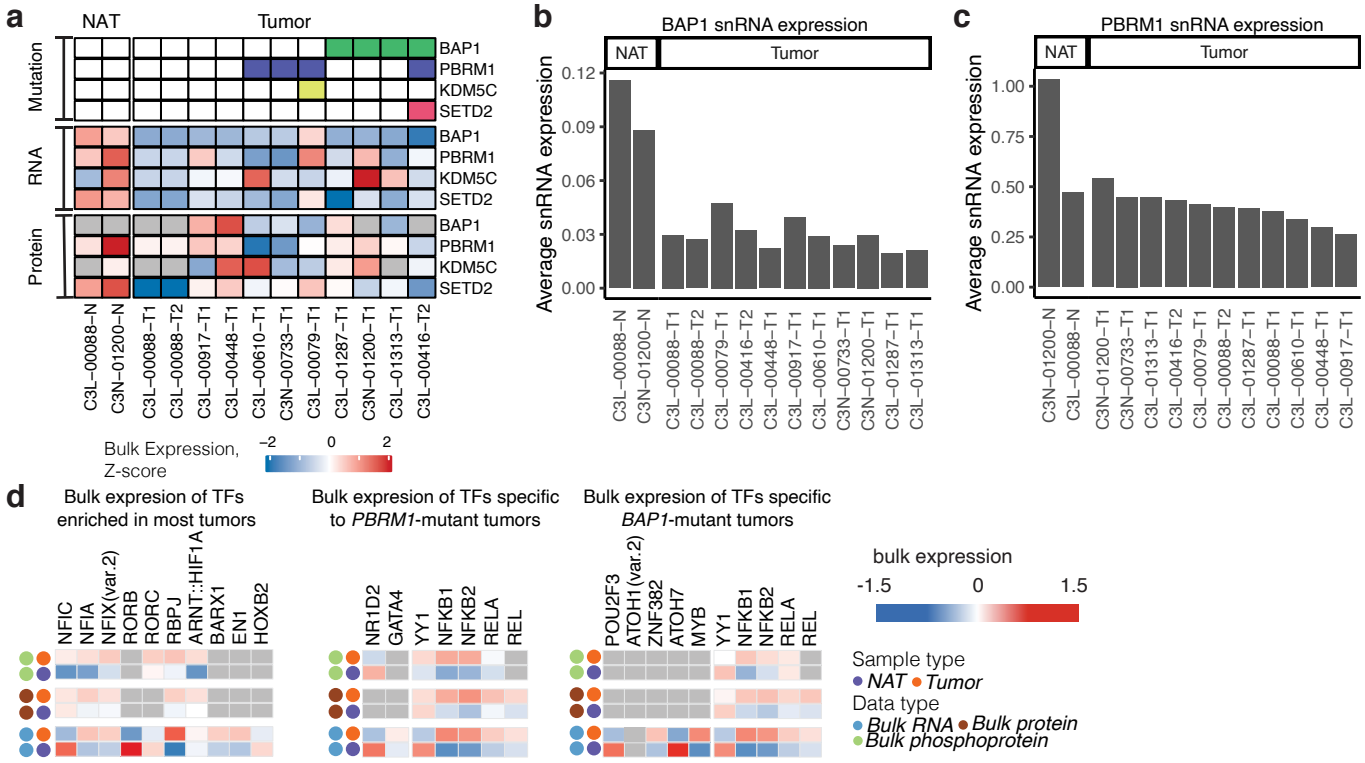

a
Immune Group
Normal cells
Tumor cells
Tumor−like epithelial cells
B−cells
CD4 CTL
CD4 T−cells
CD4 T−cells activated
CD4 T−cells naive
CD4/CD8 proliferating
CD8 CTL
CD8 CTL exhausted
CD8 T−cells preexhausted
NK cells strong
NK cells weak
Plasma
Tregs
cDC
Macrophages
Macrophages proliferating
pDC
TRM
Endothelial cells
Fibroblasts
Myofibroblasts
Cell type %
100
60
20
10
2
1
0
Bulk-based immune subtype
Metabolic immune-desert
CD8+ inflamed
VEGF immune-desert
CD8- inflamed
NAT
Non-clear-cell RCC or not in Clark et al.
b
Interacting groups
Immune:Immune
Immune:Stroma
Stroma:Stroma
Tumor:Immune
Tumor:Stroma
Tumor:Tumor
0
100
200
300
400
500
Number of interactions
c
Scaled snRNA expression
−2
−1
0
1
2
VEGFA
EGFR
IGF1R
MET
TGFA
VEGFB
HBEGF
HGF
IGF1
ANGPT1
TEK
ANGPT2
KDR
FLT1
Normal.epithelial.cells
Tumor.cells
Endothelial.cells
Fibroblasts
Myofibroblasts
Macrophages
NK.cells
CD8..T.cells
B.cells
CD4..T.cells
DC
d
Interaction strength
0
1
2
3
4
Macrophages–>Tumor cells
Transitional cells–>Transitional cells
Tumor cells–>Tumor cells

Extended Data Figure 6. Related to Figure 4.
a, Heatmap showing the cell type fraction of all the samples used in this study, grouped by the immune subtype determined by Clark et al.[13]. b, Bar plot showing the number of significant cell-cell interactions (x-axis) in each sample (y-axis). c, Heatmap showing the averaged snRNA expression of ligand and receptor genes by major cell types for the ligand-receptor pairs shown in Fig. 4. d, Heatmap showing the interaction strengths of WNT5A-FZD6 interactions across tumor samples.


a
rho = 0.84, p < 2.2e-16
Fraction of cells by cell type by sample (snATAC-seq)
Fraction of cells by cell type by sample (snRNA-seq)
b
PT, C3
PT, C4
PT, C1
PT, C2
LOH, C1
LOH, C2
UMAP_2
UMAP_1
Principal cells
Podocytes
Unknown
Distal convoluted tubule
Loop of Henle (LOH)
Intercalated cells
Proximal tubule (PT)
c
Motif score
NFIX(var.2)
NFIC
NFIA
RORB
RORC
RBPJ
BARX1
HOXB2
EN1
ARNT::HIF1A
HNF1B
NR2F1(var.2)
RXRG
RXRB
PPARD
NR1H2::RXRA
THRB
HNF4A
HNF4G
HNF4A(var.2)
d
bulkATAC-seq motif significance z-score (KIRC among pan-cancer, Corces et. al.)
snATAC-seq motif z-score (tumor cells)
HNF1A
HNF1B
HOXB4
PAX1
PAX9
POU6F1
HNF4G
HNF4A
TEAD3
TEAD4
TEAD1
RXRG
RXRB
RORB
NR3C1
NR1I3
NFIC
THRB
NR3C2
NFIA
NFIX
NFIB
PPARD
CLOCK
MXI1
TP53
TCF4
HIF1A
NEUROD2
NEUROG1
e
DEGs
TF motifs
HNF4G
HNF4A
NR2F1(var.2)
THRB
RXRB
RXRG
NR1H2::RXRA
PPARD
HNF4A(var.2)
HNF1B
Motif-DEG Association Legend
See Fig. 5 legend
PCK1
PLXNB1
SLC5A2
FGFR3
NSMF
SLC47A2
f
Endothelial-cell-specific
Fibroblasts
Endothelial cells
EHF
ELF1
ELF3
ELK4
ERF
ERG
ETS1
ETS2
ETV1
ETV2
ETV4
ETV5
ETV6
FLI1
GABPA
IKZF1
SOX8
SOX9
SPIB
ZBTB7A
Fibroblast-specific
Fibroblasts
Endothelial cells
ATF2
ATF7
CREB1
CREB3L4(var.2)
EBF1
EBF3
EGR1
EGR3
FOSL2::JUNB(var.2)
JUNB(var.2)
MEF2A
MEF2B
MEF2C
MEF2D
NFATC2
NFATC3
NR3C2
RBPJ
SRF
ZNF263
Motif Score

Extended Data Figure 7. Related to Figure 5.

a, Scatter plot showing the correlation between the fraction of cell types by sample estimated by snRNA-seq (x-axis) and snATAC-seq (y-axis). b, UMAP visualization of the nephron epithelial cells from NAT samples by snATAC-seq (n = 15,856). c, Heatmap illustrating the motif scores of top motifs consistently more (upper panel) and less (bottom panel) accessible in tumor cells compared to PT cells. d, Scatter plot showing the comparisons between the motif z-scores of the tumor cells by snATAC-seq (x-axis) vs. the motif significance z-score from the bulk ATAC-seq study[12] (y-axis). The blue dots denote the TF motifs that were specific to ccRCC compared to other cancer types by the bulk ATAC-seq study that also had high z-score in ccRCC cells compared to other cells by snATAC-seq. The purple dots denote the TF motifs specific to the snATAC-seq study but so significant for ccRCC in the bulk ATAC-seq study. e, Heatmap showing the associations between TF motifs (y-axis) with decreased accessibility in tumor cells and DEGs (x-axis) down-regulated in tumor cells compared to PT cells. The bottom panel shows the average expression of snRNA and bulk mRNA, protein and phosphoprotein of the DEGs in the tumors and NAT samples. The samples used for the bulk average expression data include 110 tumors and 75 NATs in Clark et al.[13]. f, Heatmap showing motif scores of the TF-motifs enriched in endothelial cells (upper panel) and fibroblasts (bottom panel).


a
NAT
Tumor
Mutation
RNA
Protein
BAP1
PBRM1
KDM5C
SETD2
C3L–00088–N
C3N–01200–N
C3L–00088–T1
C3L–00088–T2
C3L–00917–T1
C3L–00448–T1
C3L–00610–T1
C3N–00733–T1
C3L–00079–T1
C3L–01287–T1
C3N–01200–T1
C3L–01313–T1
C3L–00416–T2
Bulk Expression, Z-score
−2
0
2
b
BAP1 snRNA expression
NAT
Tumor
Average snRNA expression
0.12
0.09
0.06
0.03
0.00
C3L–00088–N
C3N–01200–N
C3L–00088–T1
C3L–00088–T2
C3L–00079–T1
C3L–00416–T2
C3L–00448–T1
C3L–00917–T1
C3L–00610–T1
C3N–00733–T1
C3N–01200–T1
C3L–01287–T1
C3L–01313–T1
c
PBRM1 snRNA expression
NAT
Tumor
Average snRNA expression
1.00
0.75
0.50
0.25
0.00
C3N–01200–N
C3L–00088–N
C3N–01200–T1
C3N–00733–T1
C3L–01313–T1
C3L–00416–T2
C3L–00079–T1
C3L–00088–T2
C3L–01287–T1
C3L–00088–T1
C3L–00610–T1
C3L–00448–T1
C3L–00917–T1
d
Bulk expresion of TFs enriched in most tumors
NFIC
NFIA
NFIX(var.2)
RORB
RORC
RBPJ
ARNT::HIF1A
BARX1
EN1
HOXB2
Bulk expresion of TFs specific to PBRM1-mutant tumors
NR1D2
GATA4
YY1
NFKB1
NFKB2
RELA
REL
Bulk expresion of TFs specific BAP1-mutant tumors
POU2F3
ATOH1(var.2)
ZNF382
ATOH7
MYB
YY1
NFKB1
NFKB2
RELA
REL
bulk expression
-1.5
0
1.5
Sample type
NAT
Tumor
Data type
Bulk RNA
Bulk protein
Bulk phosphoprotein

Extended Data Figure 8. Related to Figure 6.
a, Heatmap showing the bulk RNA and protein expression of BAP1, PBRM1, KDM5C, and SETD2 for snATAC-seq samples, annotated by mutation status of the same genes. b-c, Bar plots showing the snRNA expression of b, *BAP1* and c, *PBRM1* in PT cells of two NATs and in tumor cells of 11 tumors with snATAC-seq data. d, Heatmaps showing the average bulk RNA (light blue dot), protein (brown dot) and phosphoprotein (green dot) levels of the TFs enriched in tumor cells (x-axis) in tumor (orange dot) and NAT (deep blue dot) samples. Leftmost panel shows the expression of TFs enriched in most tumor samples, and the samples include 110 tumors and 75 NATs in Clark et al.[13]. The middle panel shows TFs specific to *PBRM1*-mutant tumors, and the samples include 42 *PBRM1*-mutant tumors and corresponding NATs. The rightmost panel shows TFs specific to *BAP1*-mutant tumors, and the samples include 15 *BAP1*-mutant tumors and corresponding NATs.