

Supplemental Table 1. Sample characteristics and sequencing metadata.**RNA-Seq**

GEO Accession	Patient ID/cell line	Sample type	Sample ID	Gender	Age	Input cell number	RIN	Total mapped reads (2x76bp)
GSM3290917	HIM13	Human_kidney_tubule	DS33394	F	80	2,000,000	10	244,151,932
GSM3290918	HIM13	Human_renal_cell_carcinoma	DS33395	F	80	500,000	9.9	257,438,823
GSM3290919	HIM15	Human_kidney_tubule	DS37923	M	62	2,000,000	9.9	124,435,785
GSM3290920	HIM15	Human_kidney_tubule	DS37924	M	62	2,000,000	9.9	104,505,925
GSM3290921	HIM15	Human_renal_cell_carcinoma	DS37925	M	62	125,000	9.9	99,037,346
GSM3290922	HIM15	Human_renal_cell_carcinoma	DS37926	M	62	125,000	9.8	147,534,033
GSM3290923	HIM23	Human_kidney_tubule	DS40494	M	63	500,000	7.2	246,217,904
GSM3290924	HIM23	Human_renal_cell_carcinoma	DS40496	M	63	500,000	9.1	113,565,853
GSM3290925	786-O	Human_renal_cell_carcinoma	DS34766	M	58	4,750,000	9.8	187,589,844
GSM3290926	ACHN	Human_renal_cell_carcinoma	DS37193	M	22	2,250,000	7.4	85,933,914

DNase-seq

GEO Accession	Patient ID/cell line	Sample type	Sample ID	Gender	Age	Input cell number	SPOT	Total mapped reads (2x36bp)
GSM3291010	HIM13	Human_kidney_tubule	DS26689	F	80	9,000,000	0.5650	201,833,536
GSM3291012	HIM13	Human_kidney_tubule	DS27824	F	80	9,000,000	0.4462	211,342,862
GSM3291022	HIM13	Human_renal_cell_carcinoma	DS26693A	F	80	5,000,000	0.5721	295,184,545
GSM3291023	HIM13	Human_renal_cell_carcinoma	DS26692B	F	80	5,000,000	0.4865	39,174,471
GSM3291014	HIM15	Human_kidney_tubule	DS37969	M	62	2,000,000	0.3638	53,455,680
GSM3291015	HIM15	Human_kidney_tubule	DS37971	M	62	2,000,000	0.4493	350,791,957
GSM3291016	HIM15	Human_renal_cell_carcinoma	DS37973	M	62	300,000	0.4951	260,802,880
GSM3291017	HIM15	Human_renal_cell_carcinoma	DS37974	M	62	300,000	0.5024	55,333,228
GSM3291020	HIM23	Human_kidney_tubule	DS41160	M	63	80,000	0.3944	221,916,751
GSM3291021	HIM23	Human_kidney_tubule	DS41396	M	63	80,000	0.5052	34,333,942
GSM3291018	HIM23	Human_renal_cell_carcinoma	DS40508	M	63	80,000	0.2764	201,007,725
GSM3291019	HIM23	Human_renal_cell_carcinoma	DS40509	M	62	80,000	0.2179	55,078,250
GSM3291011	786-O	Human_renal_cell_carcinoma	DS27192	M	58	9,000,000	0.3238	214,308,625
GSM3291013	786-O	Human_renal_cell_carcinoma	DS37199	M	58	100,000	0.3742	360,466,731
GSM3291024	ACHN	Human_renal_cell_carcinoma	DS24547A	M	22	13,100,000	0.4583	261,250,750
GSM3291025	ACHN	Human_renal_cell_carcinoma	DS24471A	M	22	3,000,000	0.4256	224,800,319

Supplemental Figure 1. Characterization of primary RCC cultures and overview of DHS landscape

(a) *Karyotypes of primary RCC cultures.* Primary RCC cultures were submitted for G-band karyotyping at the University of Washington Cytogenetics Laboratory. Inferred karyotypes from 20 metaphase spreads are shown. All three patient tumors showed characteristic loss of chromosome 3p and gain of chromosome 5q.

(b) *VHL mutation status in primary tubule and RCC cultures.* Sanger sequencing of the coding regions identified an inactivating mutation in the single copy of the *VHL* gene in all three primary RCC cultures.

(c) *Chromosome arm level gains and losses are identified by DNase-seq tags.* Windowed aggregation (5Mb windows) of tags from DNase-seq datasets from the primary tubule and RCC cultures revealed arm level gains and losses, including the canonical loss of chromosome 3p and gain of chromosome 5q.

(d) *DNase I-hypersensitive sites identify predominantly distal regulatory elements.* A minority of the master list DHS derived from tubule and RCC datasets localized to promoter elements (1.9%) or lie within 5 kb of a known transcription start site (20.3%). The majority (77.8%) lay >5 kb from known transcription start sites, characteristic of distal regulatory elements such as enhancers.

(e) *Overlap of differential DHS identifies the shared regulatory landscape of RCC.* DHSs with differential accessibility in tumors vs. their matched tubule controls define the differentially accessible regulatory landscape of RCC. Pairwise comparisons of these differential DHS across patients revealed that ~35% of all differential DHS are shared among at least two patient samples.

Supplemental Figure 2. Individual renal cell carcinomas exhibit highly concordant RNA landscapes

(a) *Consistent patterns of gene expression among patients.* Comparison of expression fold change of genes (translucent blue dots) revealed largely consistent patterns of gene expression among patient tumor samples. Genes that typify the HIF transcriptional response (e.g. *CA9*, *VEGFA*, *EGLN3*) were upregulated and some genes associated with normal tubular function (e.g. *CDH1*, *ANPEP*) were downregulated in all three tumor samples compared to their normal tubule controls. Hatched lines represent 1.5x fold change on the log-scale axis.

(b) *Gene ontology enrichment.* Ranked list GOrilla enrichment analysis (rank in boldface) identified both canonical (e.g. hypoxia response, lipid metabolic process, chromatin organization) and unexpected gene ontologies (e.g. complement activation) that were differentially regulated in renal cell carcinoma compared to tubules.

Supplemental Figure 3. Expression of *PSORS1C3* in various TCGA tumors and matched normal tissues. The mRNA expression levels of the *PSORS1C3* gene in several non-germ cell tumors was compared to their matched normal tissue controls. The ends of the bar plots represent the 25th and 75th quartiles with whiskers representing 1.5x inter-quartile range (10% outlier trim applied for clarity).

Supplemental Figure 4. An adult kidney-specific DHS encodes an alternate promoter for *POU5F1*

The alternate promoter for *POU5F1* in RCC maps to a DHS that is only present in adult kidney derived tubules (RPTEC, renal proximal tubule epithelial cells), primary cultures (HRCE, human renal cortical epithelial cells; HRE, human renal epithelial cells) or tumors (786-O, ACHN). This DHS is not present in fetal kidney, embryonic stem cells, non-epithelial kidney cells (e.g. glomerular endothelial cells, HRGEC) or a variety of ontologically diverse cells. All tracks are shown at same scale/settings as those shown in Fig. 5.

Supplemental Figure 5. Numerous expressed sequence tags (ESTs) in the *POU5F1-PSORS1C3* genomic locus (hg19 chr6:31,125,253-31,156,354)

Known transcript isoforms for *POU5F1*, *PSORS1C3* and *TCF19* are shown at the top in blue gene models. Annotated ESTs in GenBank are shown below in black gene models.

Supplemental Figure 6. Overview of the *POU5F1B* genomic locus (hg19 chr8:128,420,724-128,436,573)

POU5F1B was expressed in human ESCs, but not in the primary tubule and RCC cultures described in this study. There were no DHSs in this genomic interval, and there is negligible binding of HIF components. Histone modifications typical of active transcription were also not present. All tracks are shown at same scale/settings as those shown in Fig. 5.

Supplemental Figure 7. Expression of *POU5F1* in RCC as a function of known metastasis status and pathologic stage. Each point represents the expression value of a single patient's tumor in the indicated subgroup. Black diamond indicates mean value for the indicated subgroup in both panels.

(a) Expression of *POU5F1* as a function of known metastasis status.

(b) Expression of *POU5F1* as a function of pathologic stage at the time of diagnosis. P-value = 0.037 by 1-factor ANOVA.

Supplemental Figure 1

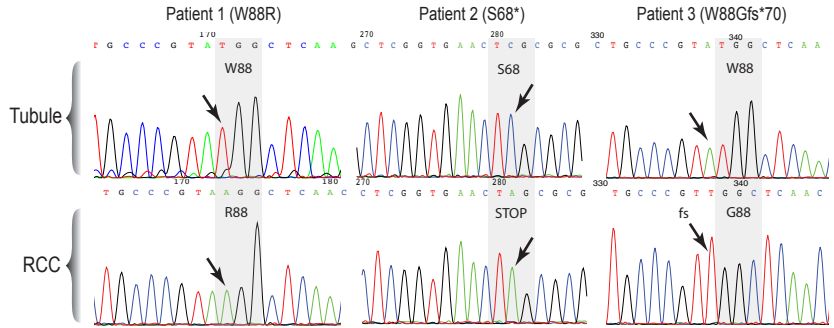
a Karyotypes of primary RCC cultures

Patient 1
46,XX,der(3)t(3;5)(p13;q22)[14]/
45,sl,der(8;9)(q10;q10)[6]

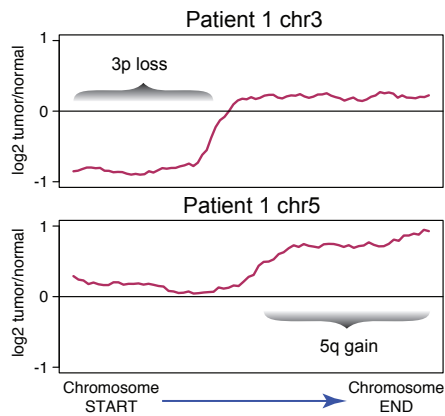
Patient 2
46,XY,der(5)t(5;15)(p15.1;q11.2),add(13)(p12),-15,+r[3]/
46,XY,del(5)(p11),add(13)(p12),der(14;?)(p11.1;?)(q11.2),-15,+r[2]/
43-45,XY,der(3)t(3;5)(p13;q31),-9,-13,-14,+1-2mar[cp2]/
47,XY,+10[2]/46,XY[11]

Patient 3
44,XY,add(1)(p36.1),der(3)t(3;5)(p21;q22),add(5)(p13),
der(8;18)(q10;p10),-14[18]/82-86,slx2[2]

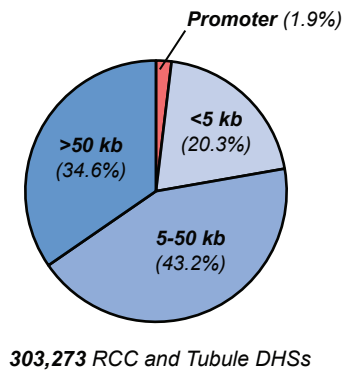
b Inactivating mutations in *VHL* gene in primary RCC cultures



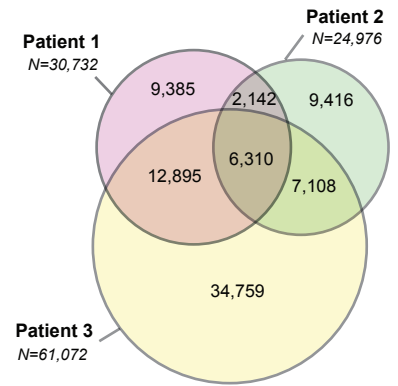
c Windowed aggregation of DNase-seq tags identifies chromosome arm level gains/losses



d Distance of RCC and tubule DHSs to nearest transcription start site

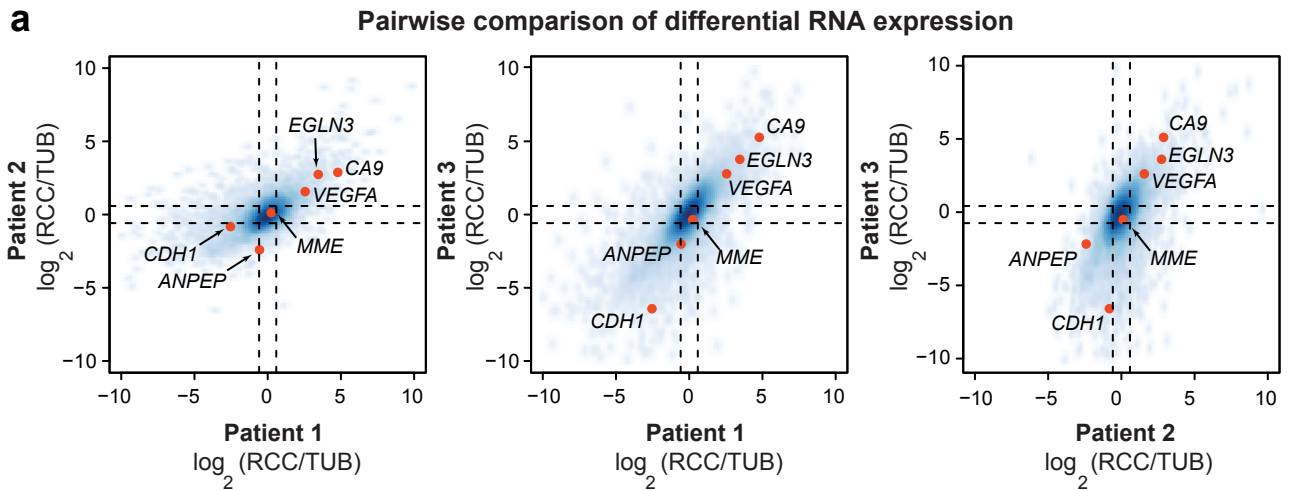


e Pairwise overlap of differential DHSs

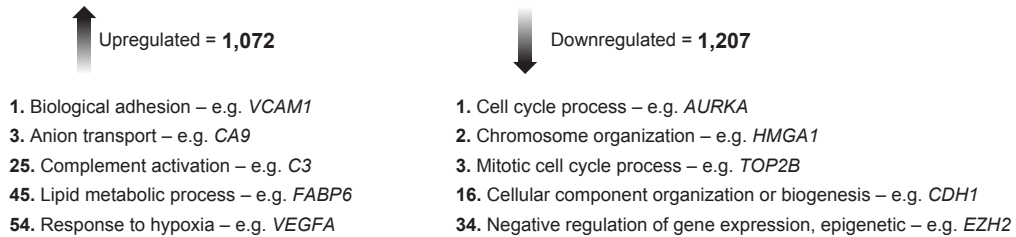


35% of dDHSs (28,455/82,015) are shared by at least 2 patients

Supplemental Figure 2

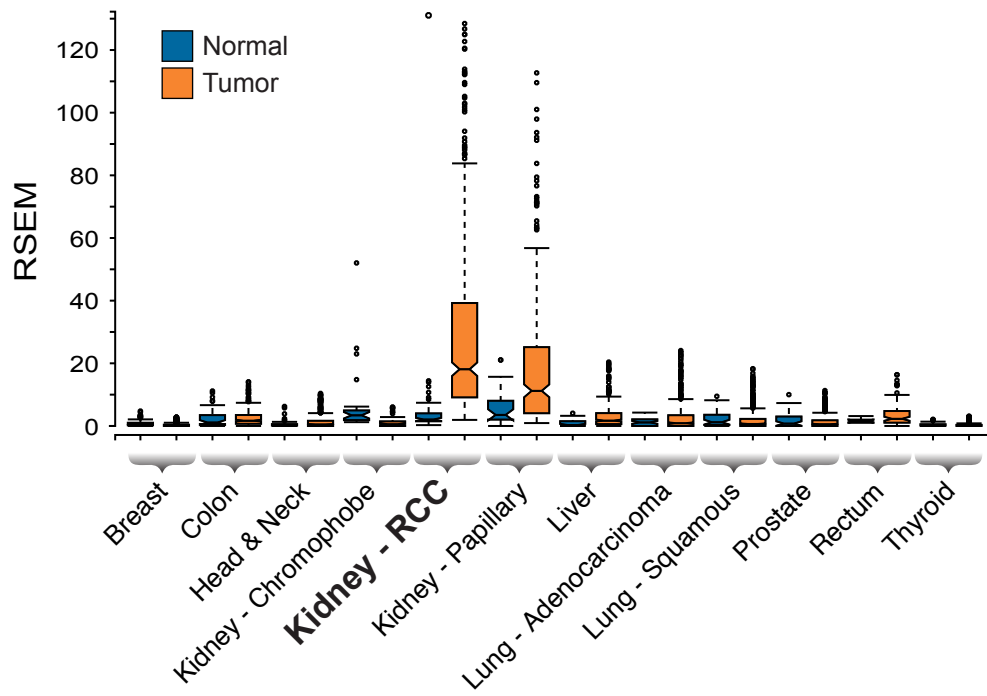


b Gene ontology enrichment of genes changing >1.5x in all 3 patients



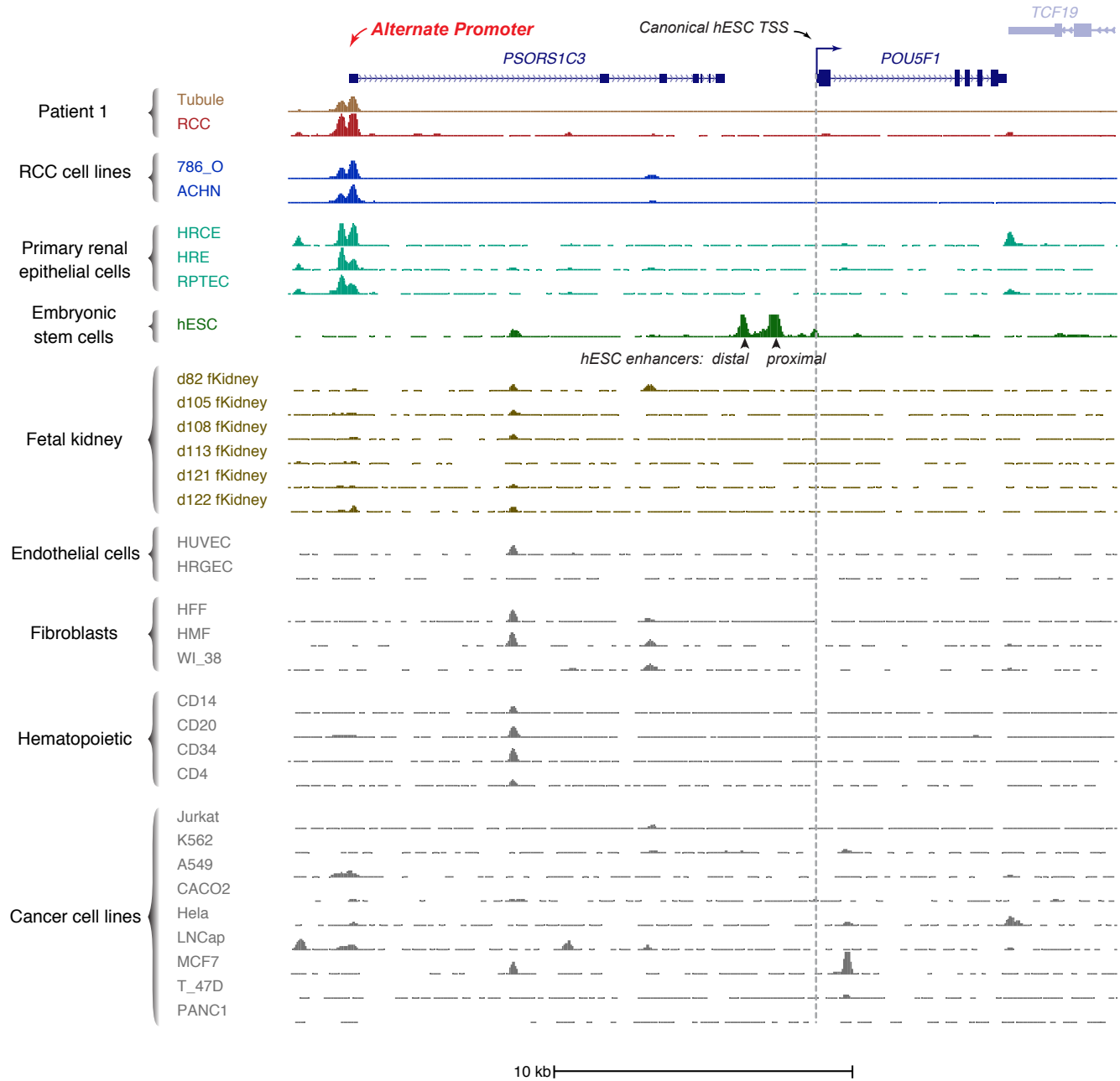
Supplemental Figure 3

Expression of *PSORS1C3* in various TCGA tumors and matched normal tissues



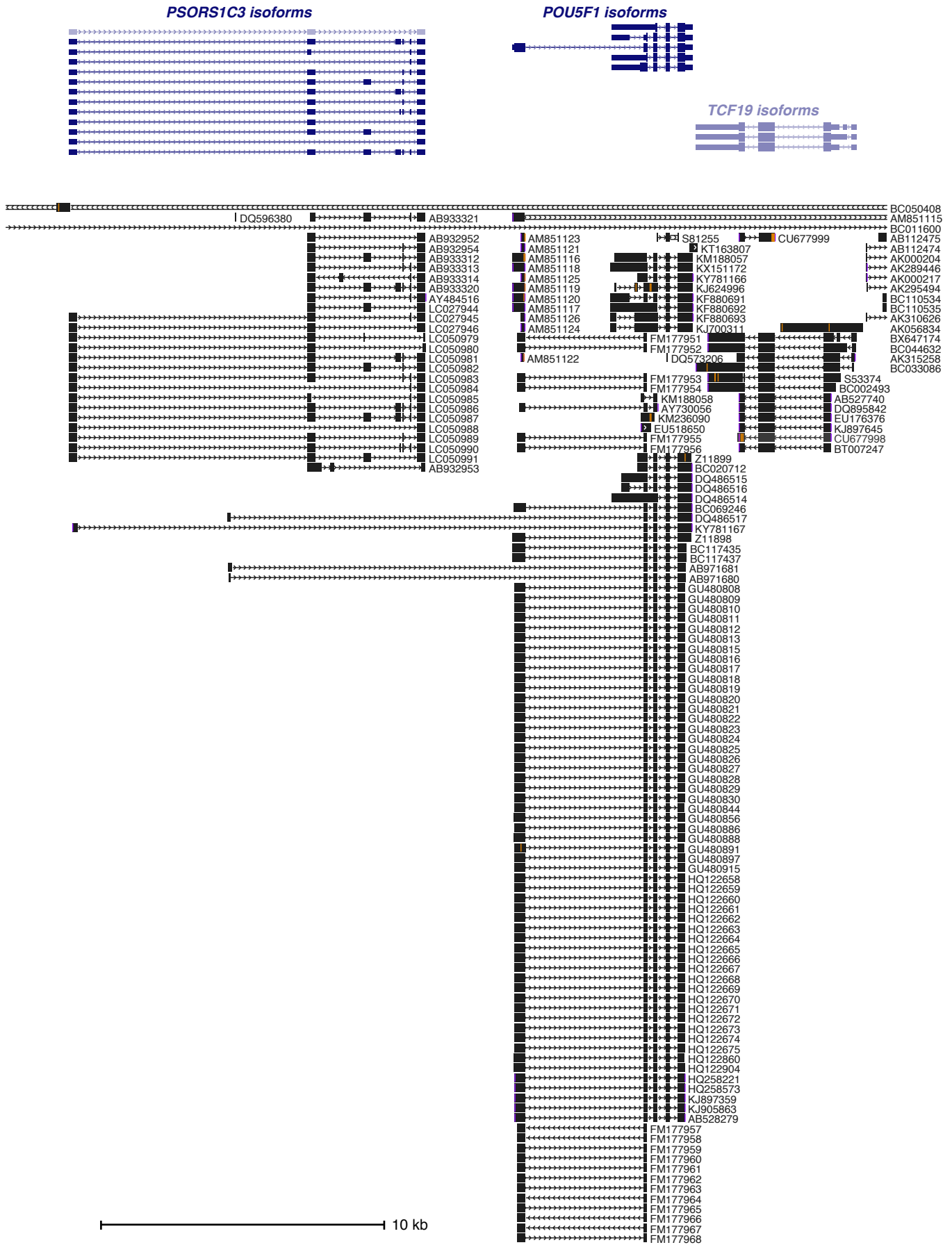
Supplemental Figure 4

The alternate *POU5F1* promoter maps to an adult kidney epithelial-specific DHS



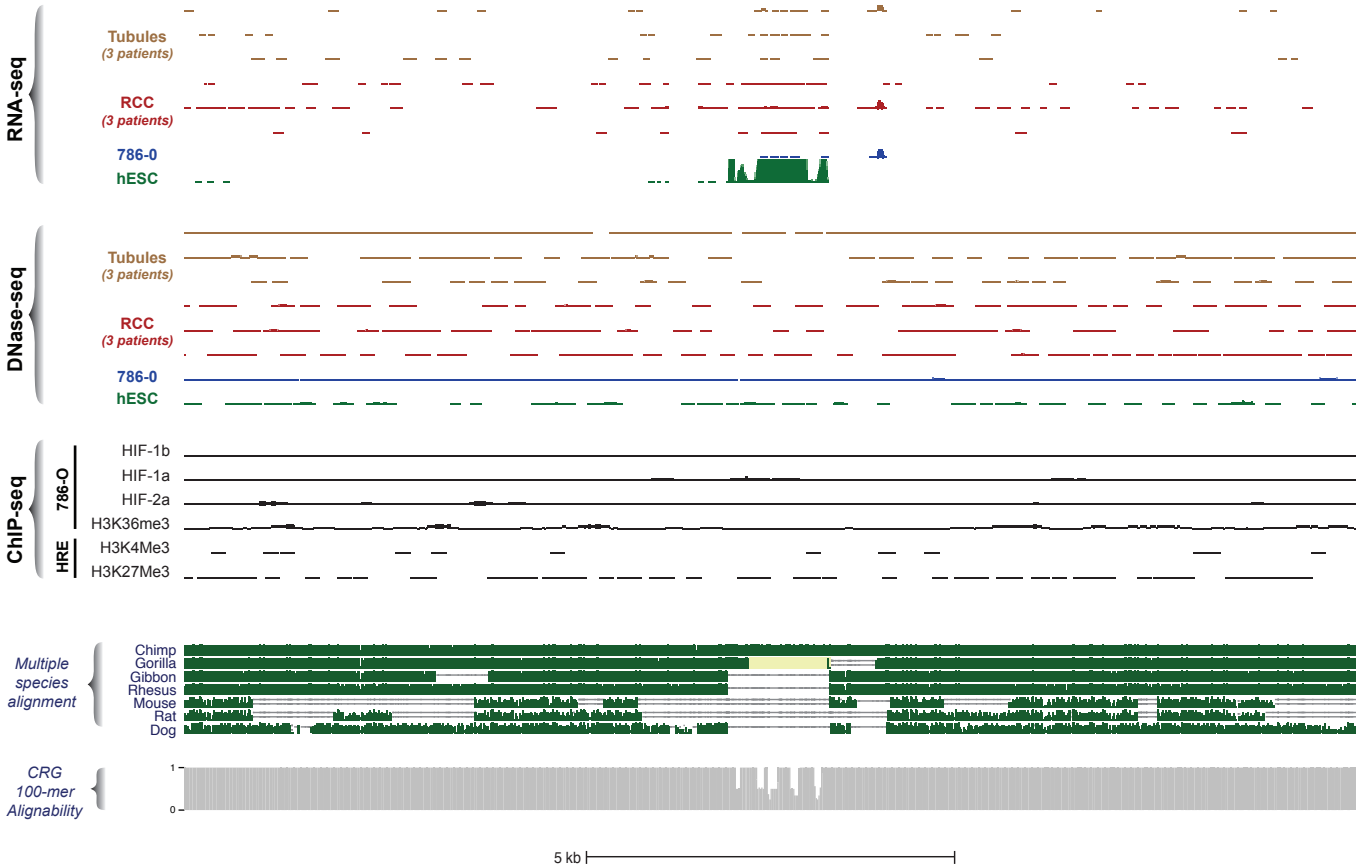
Supplemental Figure 5

Numerous expressed sequence tags (ESTs) in the *POU5F1-PSORS1C3* locus



Supplemental Figure 6

POU5F1B
▶▶▶▶▶▶▶▶▶▶



Supplemental Figure 7

