

Supplementary Information

The Fibroblast Epigenome Underlies SS18::SSX-Mediated Transformation in Synovial Sarcoma

Lesley A. Hill^{*1,2}, R. Wilder Scott^{*1,2}, Lauren A. Martin^{1,2}, Martin Arostegui^{1,2}, George Davenport^{1,2}, Marcos Vemon^{1,2}, Jakob Hofvander³, Xue Qi Wang⁴, Jinxiu Li^{5,6}, Torsten O. Nielsen⁴, Kevin B. Jones^{5,6}, Martin Hirst⁷, and T. Michael Underhill^{1,2#}

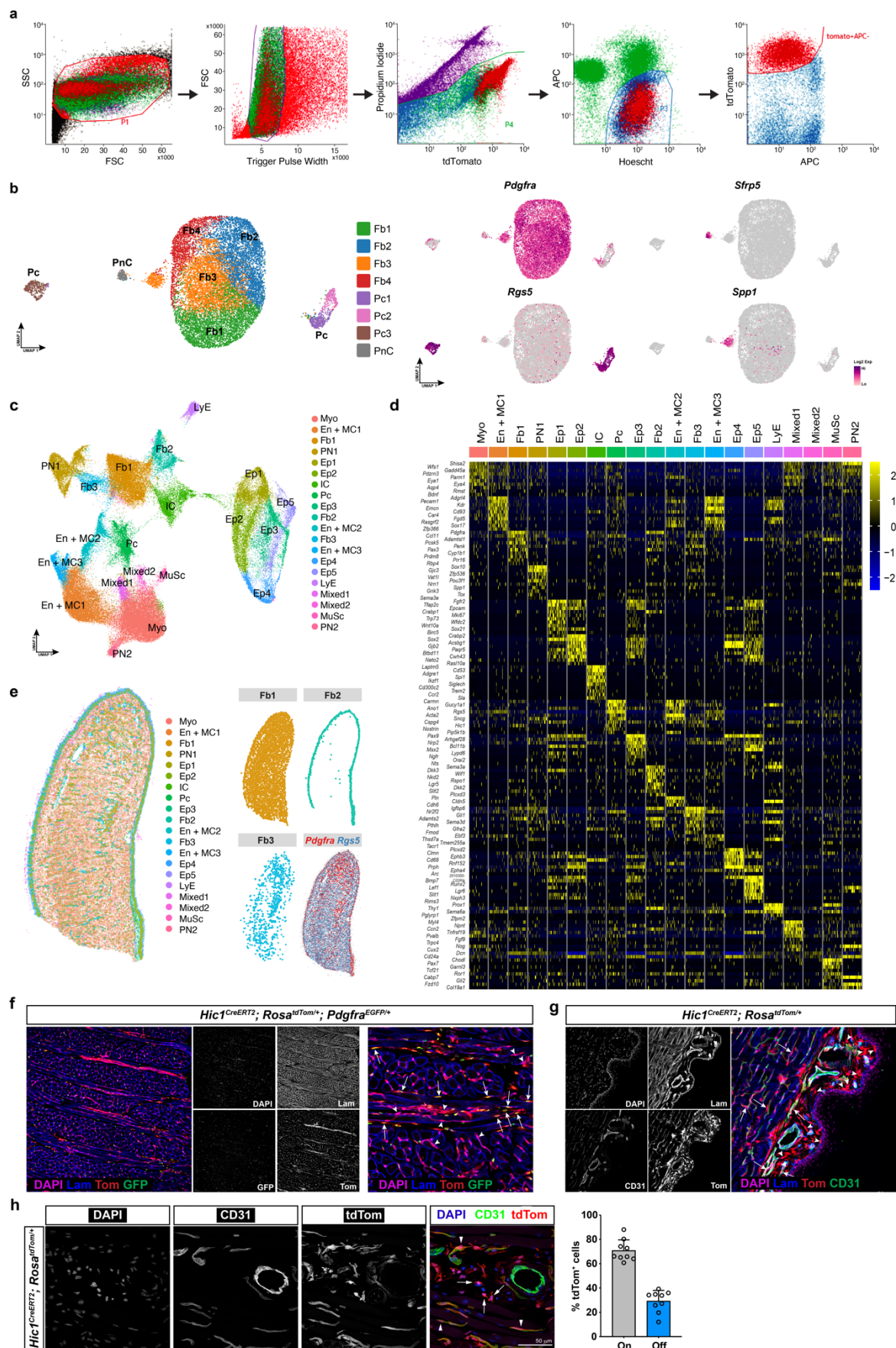
¹Biomedical Research Centre, ²Department of Cellular and Physiological Sciences, 2222 Health Sciences Mall, University of British Columbia, Vancouver, BC, Canada, V6T 1Z3; ³Division of Clinical Genetics, Department of Laboratory Medicine, Lund University, Lund, Sweden; ⁴Department of Pathology, University of British Columbia, Vancouver, British Columbia, Canada; ⁵Department of Orthopaedics, University of Utah, Salt Lake City, Utah, United States of America; ⁶Department of Oncological Sciences, Huntsman Cancer Institute, Salt Lake City, Utah, United States of America; ⁷Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada.

*These authors contributed equally.

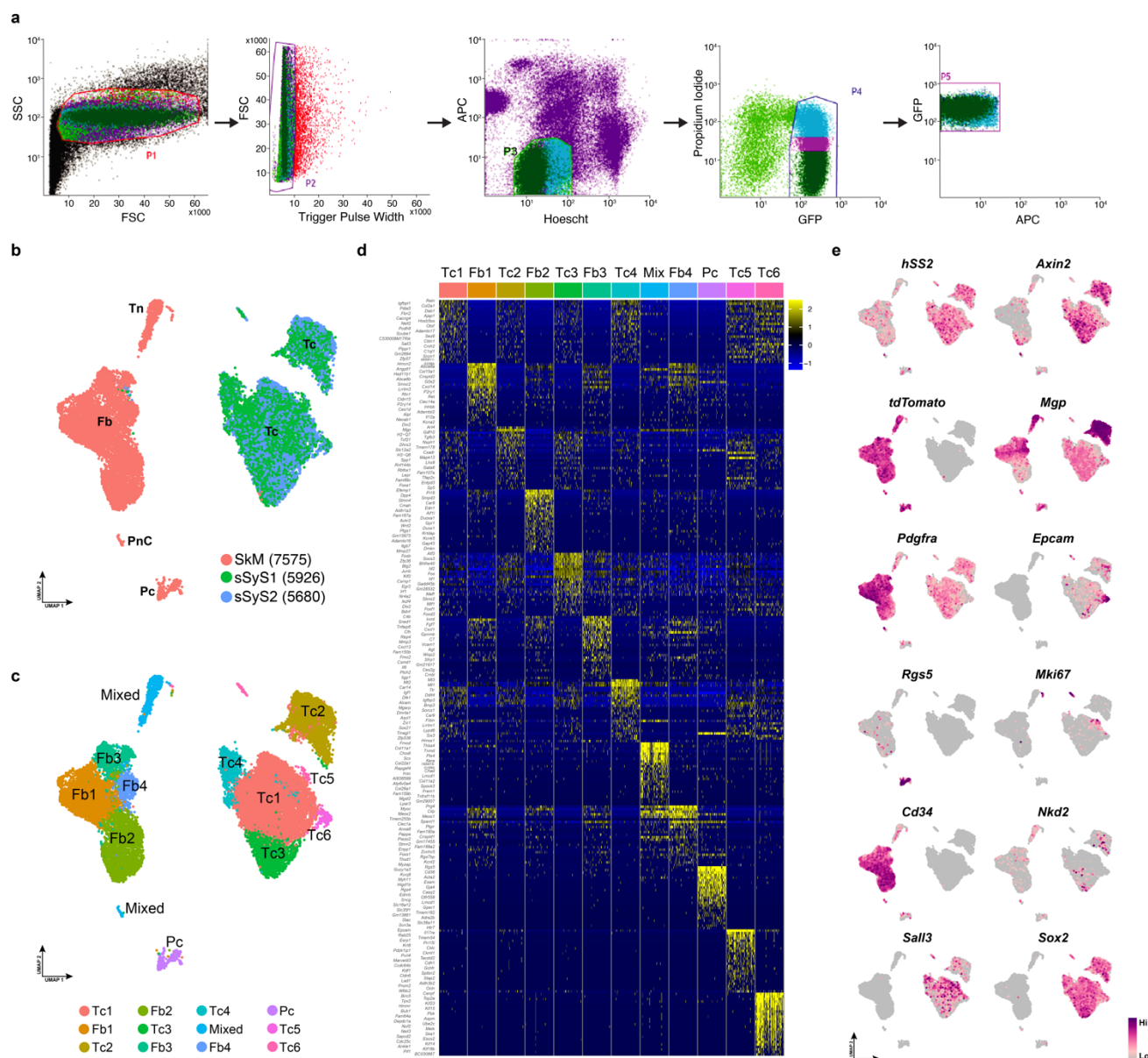
Corresponding Author: tunderhi@brc.ubc.ca.

Supplementary Figure 1: Expression of SS18::SSX in *Hic1*⁺ MSCs induces tumours with features of SyS.

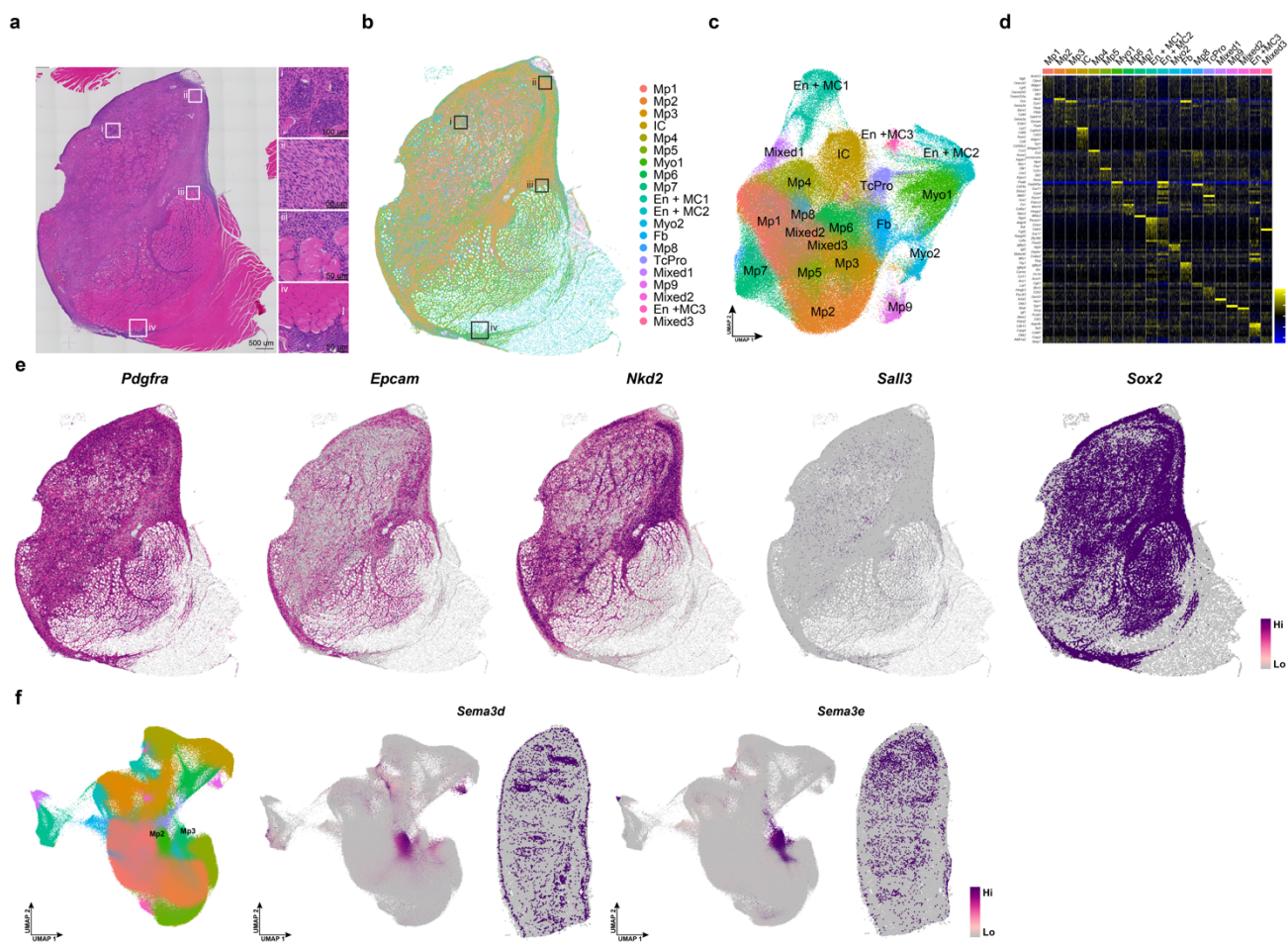
a, Kaplan Meier survival curves for the indicated genotypes, separated based on sex. **b**, Kaplan Meier survival curve for mice in which SS18::SSX was ubiquitously expressed using UBC-CreERT2, following TAM administration at 8 weeks of age. **c**, Whole mount images of GFP⁺ tumours (*Hic1*^{CreERT2}; *Rosa*^{hSS2}) at various anatomical sites. T-L, thoracic-lumbar vertebral column junction. **d-g**, Immunofluorescence staining of sections from end-point tongue (**d**) and stifle tumours (**e**) for MUC1 (**d**; broken lines indicate areas of pseudogland formation). E, epithelioid; F, fascia; M, muscle; T, tumour. **f**, SS18::SSX (dotted lines demarcate tumour margin) and TLE1 in tongue and stifle tumours. **g**, Corresponding immunostaining of control tongues (*Hic1*^{CreERT2}; *Rosa*^{tdTomato/+}). **h**, Schematic overview of the experimental plan for evaluating cell proliferation using EdU incorporation. **i-k**, Detection of EdU⁺ cells in control tongue sections (**i**; white box region is enlarged in right panel with a rare single EdU⁺ tdTom⁺ cell highlighted by arrow) and within GFP⁺ tongue tumours (**j**) collected at end-point (~16 weeks post-TAM), counterstained with DAPI and the indicated antibodies. **k**, Enumeration of EdU⁺ cells in control (n=3) and tumour-bearing mice (n=6), as a percentage of reporter⁺ cells. Data are presented as mean ± SD with each dot representing an individual section (n = 3 sections/mouse). Statistical comparison was performed using a two-tailed *t*-test. **l**, Correlation matrix heatmap of Euclidean distance between tumour samples (hSS2/hSS2) and controls, collected from the tumour origin site (tongue and stifle region), in *Hic1*^{CreERT2} (*Hic1*-Cre), *Myf5*-Cre and TAT-Cre mice, based on RNA-seq gene expression. **m**, Principal component analysis plot of the variance explained by the first two components across the samples from (**l**), colored by sample group. In **d-g**, **i-j**, experiments were repeated independently at least 3 times with similar results. Representative images/data from one experiment are shown. Source data are provided in the Source Data file.



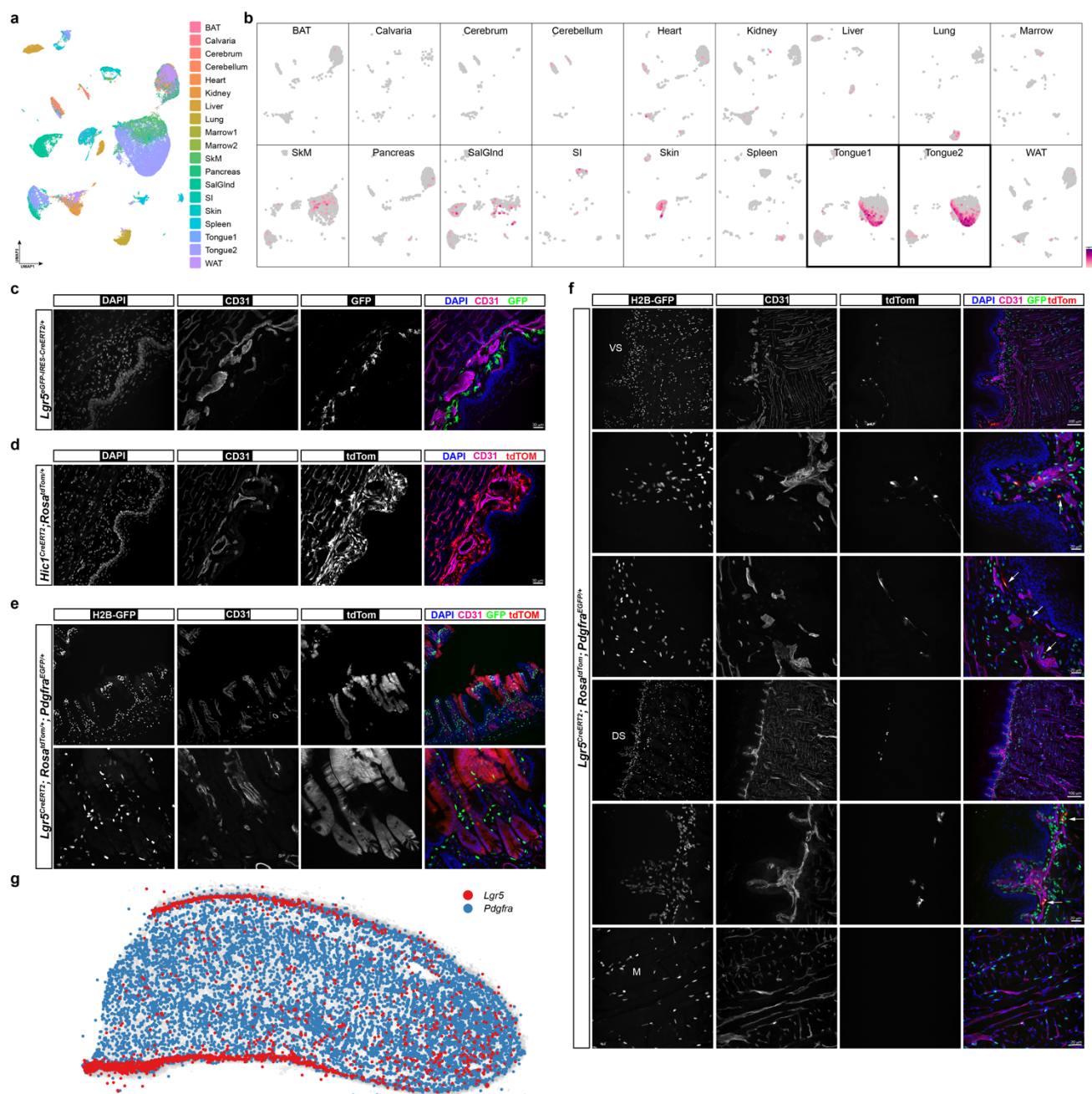
Supplementary Figure 2: Cellular and molecular characterization of *Hic1*⁺ MSCs in the tongue. **a**, Flow cytometry strategy used to sort tdTom⁺ MSC populations from control tongues for scRNA- and ATAC-seq analyses. **b**, UMAP of scRNA-seq samples (n=2) generated from *Hic1*⁺ MSCs (*Hic1*; tdTom⁺) colored by cluster identity. K-means (k=8) was used for clustering, and cell types are annotated based on transcripts representative of different lineages: Fb, fibroblast; Pc, pericyte; PnC, perineural. **c**, UMAP colored by cluster identity derived from spatial transcriptomic analysis, annotated by marker analysis in (d). **d**, Heatmap of top 10 defining transcripts per cluster, with their cellular identity defined as: En + MC, endothelial and mural; Ep, epithelial; Fb, fibroblast; IC, immune; LyE, lymphatic endothelium; MuSc, muscle satellite cell; Myo, myofibre; Pc, pericyte; PN, peripheral nerves. **e**, Mapping of identified clusters to spatial domains, with the distribution of Fb cell types highlighted on the right. Spatial distributions of *Pdgfra* and *Rgs5* transcripts are also shown. **f-g**, Representative images of *Hic1*^{CreERT2}; *Rosa*^{tdTom/+}; *Pdgfra*^{EGFP/+} cells in the medial (f) and ventral (g) compartments of the tongue. Arrows identify tdTom⁺ (Tom) EGFP⁺ (GFP) cells; arrowheads denote Tom⁺ GFP⁻ cells. Nuclei, basement membrane and vasculature counterstained with DAPI, anti-Laminin and anti-CD31, respectively. **h**, Representative images of *Hic1*^{CreERT2}; *Rosa*^{tdTom/+} cells within the tongue, and enumeration of their proximity to CD31-positive vessels (n=3, 3 independent sections from each biological replicate). Source data are provided in the Source Data file. In **f-h**, experiments were repeated independently at least 3 times with similar results. Representative images from one experiment are shown.



Supplementary Figure 3: Integrated transcriptomic profiling of stifle tumours. **a**, Flow cytometry strategy used to sort GFP⁺ tumor cell populations for scRNA- and ATAC-seq analyses. **b-c**, UMAPs of scRNA-seq data aggregated from the indicated samples, colored by sample ID (b) and cluster ID (c). SkM, tdTom⁺ cells (*Hic1*; *tdTom*⁺) from hind limb skeletal muscle and overlying fascia; sSyS, stifle sarcomas. Cell types are also indicated: Pc, pericyte; PnC, perineural cell; Tc, tumour; Tn, *Scx*⁺ tenogenic. **d**, Heatmap of scRNA-seq data showing the top 25 enriched genes in the 12 different clusters in (c). **e**, UMAPs colored by expression level of various cluster-enriched transcripts.

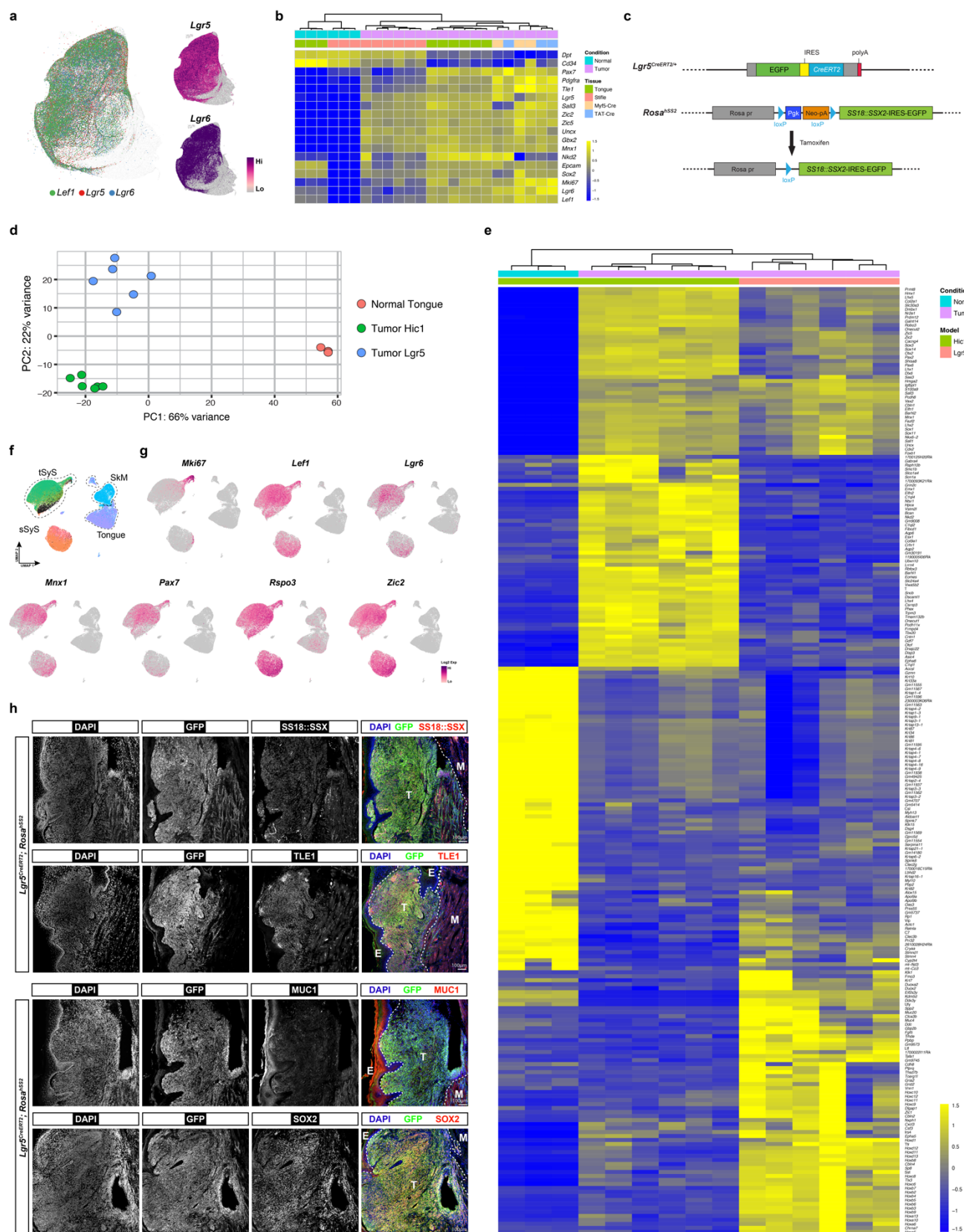


Supplementary Figure 4: Spatial transcriptomic profiling of tumours. **a**, H&E whole-mount section of a representative stifle tumour (sSyS3 in Figure 3) processed for spatial transcriptomics, with insets (i, ii, iii, iv). **b**, Cell clusters identified based on transcript expression, mapped to spatial domains on the image in (a), with their assignment to different tumour phenotypes and cell types specified; IC, immune; Myo, myofibre; En + MC, endothelial and mural; Fb, fibroblast. **c**, UMAP of cells identified in (b) colored by cluster identity. **d**, Heatmap of the top 10 cluster markers for the clusters in (c-d). **e**, Spatial localization of select indicated transcripts in stifle tumours colored by expression level. **f**, UMAPs (integrated tongue and stifle tumors) and spatial localization (tSyS2) of select transcripts in tongue tumours colored by expression level. In **a-e**, four independent stifle tumors (sSyS1-4) were processed for Xenium analyses with similar results (see Figure 3).

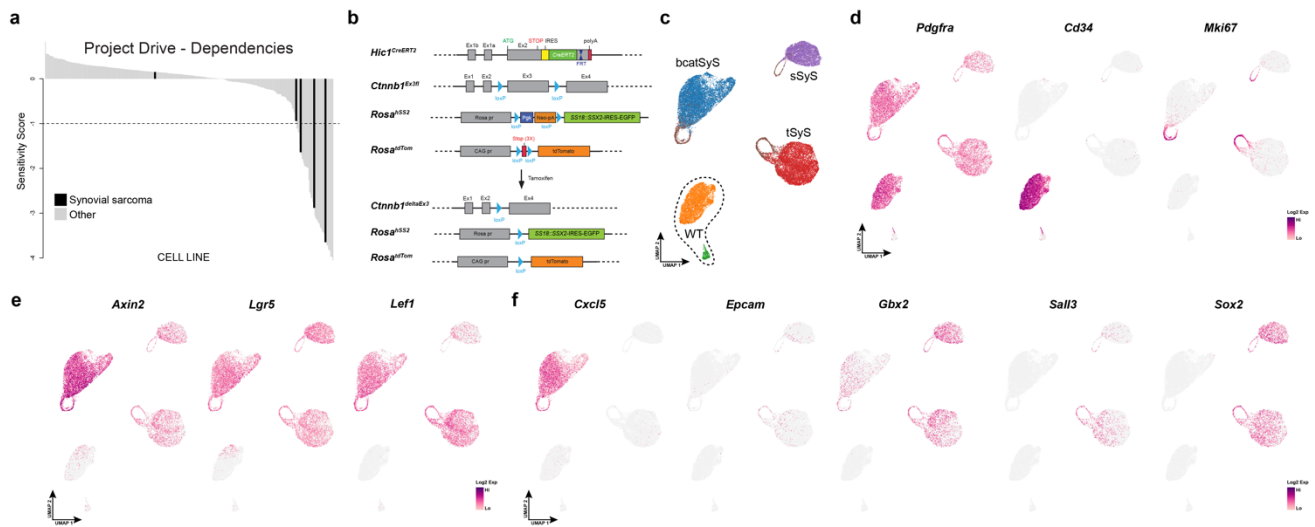


Supplementary Figure 5: Characterization of *Lgr5*⁺ fibroblasts in the tongue. **a**, UMAP of aggregated scRNA-seq data colored by sample ID, comprising *Hic1*⁺; *tdTom*⁺ MSCs isolated and enriched from the indicated tissues of 10-week old mice, 10 days post-TAM induction. BAT, brown adipose tissue; Marrow, bone marrow from fore- and hindlimb bones; SalGnd, salivary gland; SI, small intestine; WAT, white adipose tissue. **b**, UMAP from (a), split by tissue type, colored by *Lgr5* expression level. **c-f**, Representative images of: **c**, Immunodetection of GFP and CD31 in the ventral lamina propria of *Lgr5^{CreERT2-GFP}* mice. **d**, Detection of the tdTom⁺ population (*Hic1^{CreERT2}; Rosa^{tdTom/+}* mice) within the ventral region of the tongue. **e**, Lineage tracing of *Lgr5* progeny in the small intestine using the *Rosa^{tdTom/+}* reporter gene in the indicated genetic background. Samples were collected 10 days post-TAM administration. **f**, Detection of *Lgr5* lineage traced cells (tdTom⁺)

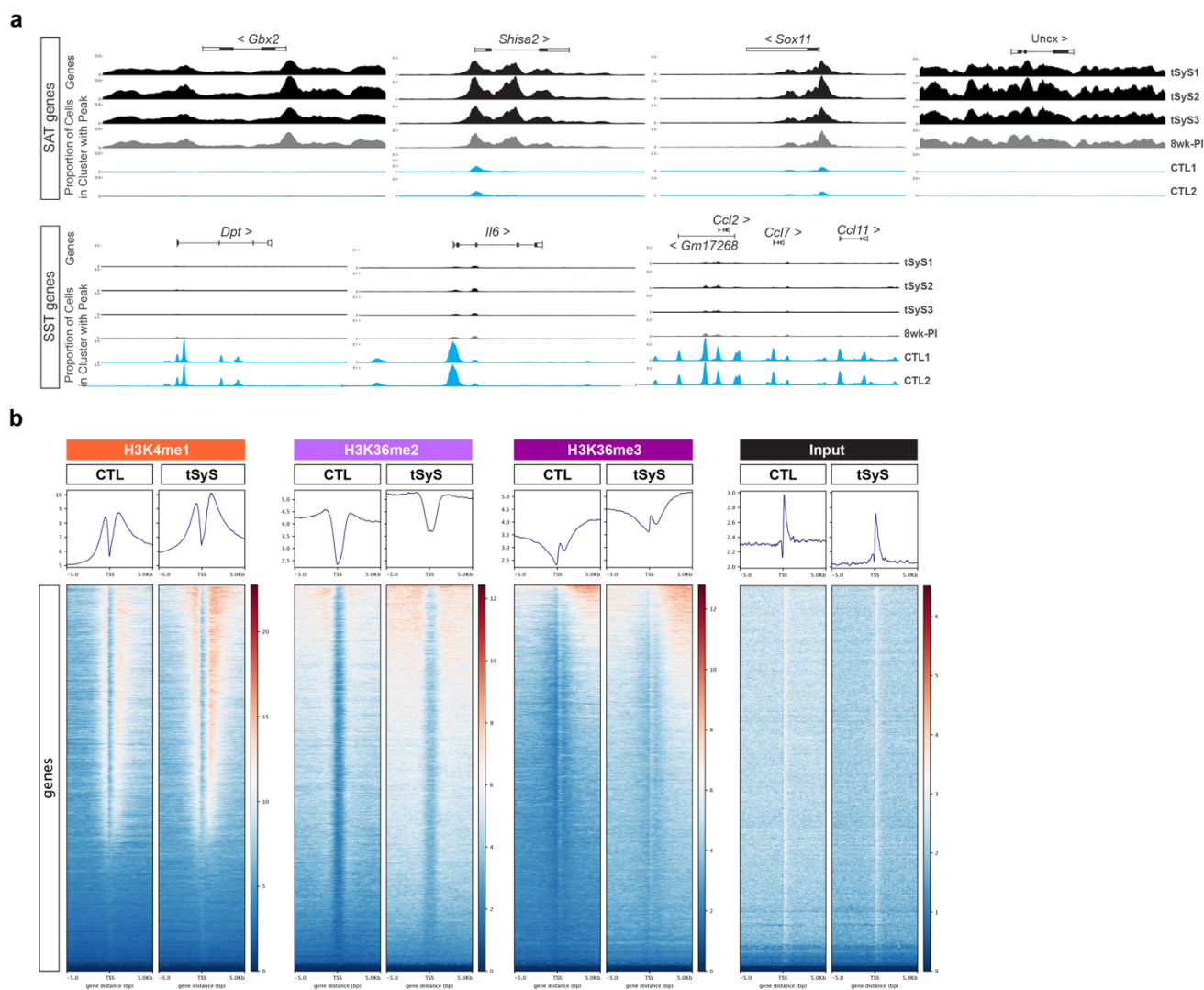
and *Pdgfra*⁺ cells (nuclear H2B-GFP⁺). Arrows denote doubly positive tdTom and nuclear GFP cells in the dorsal and ventral regions of the tongue. DS, dorsal surface; M, midline; VS, ventral surface. **g**, Distribution of *Lgr5* and *Pdgfra* transcripts in the wild-type tongue, as determined by spatial transcriptomic analysis. In **c-f**, experiments were repeated independently at least 3 times with similar results. Representative images from at least one experiment are shown.



Supplementary Figure 6: *Lgr5*⁺ MSC-derived tumours parallel *Hic1*⁺ MSC resultant SyS tumours. **a**, Spatial transcriptomic analysis of *Lef1*, *Lgr5* and *Lgr6* distribution in stifle tumours. **b**, Heatmap showing differential gene expression from a subset of transcripts, between tumour samples (hSS2/hSS2) and controls, collected from the tumour origin site (tongue and stifle region), in *Hic1*^{CreERT2} (*Hic1*-Cre), *Myf5*-Cre and TAT-Cre mice. **c**, Experimental plan for induction of SS18::SSX in *Lgr5*-expressing cells. **d**, Principal component analysis plot of the variance explained by the first two components across tumour samples (hSS2/hSS2) and controls, collected from the tumour origin site (tongue region), in *Hic1*^{CreERT2} (*Hic1*-Cre) and *Lgr5*^{CreERT2} (*Lgr5*-Cre) mice, colored by sample group. **e**, Heatmap showing differential gene expression (top 50 up and down regulated transcripts per group) between the samples in (d). **f**, scRNA-seq UMAP of cells from the indicated samples, colored by sample ID, with the *Lgr5*-tumour derived sample highlighted in black. **g**, UMAPs from (f) colored by the expression level of the indicated select transcripts. **h**, Representative images of the immunodetection of tumour-associated proteins (SS18::SSX, TLE1, MUC1 and SOX2), in relationship to GFP⁺ tumour cells, from sections of *Lgr5*^{CreERT2}; *Rosa*^{hSS2/hSS2} tumours, counterstained as indicated. Experiments were repeated independently at least 3 times with similar results.



Supplementary Figure 7: Activation of WNT/β-catenin signaling promotes a program consistent with monophasic tumours. **a**, Waterfall plot of *CTNNB1* dependencies as determined in the DRIVE project. Each black line indicates a different human synovial sarcoma cell line. **b**, Schematic of the alleles used to activate WNT/β-catenin signaling in the *Hic1*; *hSS2* background. **c-f**, UMAPs of scRNA-seq data, colored by sample ID (**c**) and by the expression level of select differentially expressed genes representative of normal fibroblasts (**d**), WNT signaling components (**e**) and mid-to-late stage transformed cells (**f**).



Supplementary Figure 8: Modification of the epigenome by SS18::SSX. a, Genome browser tracks derived from scATAC-seq, displaying the promoter sum accessibility signal around specified gene loci from the indicated samples. tSyS, endpoint (~16 weeks post TAM) tongue tumour; 8wk-PI, 8 weeks post-induction of SS18::SSX; CTL, control tongue. Tissue samples were collected at the indicated time points, and reporter⁺ cells were enriched by FACs and processed for scATAC-seq. **b**, Profile heatmaps of the indicated histone modifications, detected by ChIP-seq, around the TSS (+/- 5kB) of RefSeq genes. The gradient, blue-to-red, represents low-to-high counts in the noted regions.

Supplementary Figure 9: Broad “re-activation” of genes within the HoxA, B, C and D clusters through SS18::SSX-mediated reprogramming. **a**, Heatmaps of indicated histone modifications observed at the *HoxA*, *B*, *C* and *D* loci (TSS +/- 5kB). **b**, Individual genome track plots of indicated histone modification peaks detected by ChIP-seq, along with transcript abundance generated from RNA-seq (+ strand (F) and – strand (R) tracks), within the HoxA, B, C and D clusters. Annotation track denotes exon and intron structure, along with coding information. CTL, control tongue; tSyS, endpoint (~16 weeks post TAM) tongue tumour.

Supplementary Table 1. Genotyping Primers

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')
<i>Cre</i>	GCG GTC TGG CAG TAA AAA CTA TC	GTG AAA CAG CAT TGC TGT CAC TT
<i>GFP</i>	AAG TTC ATC TGC ACC ACC G	TCC TTG AAG AAG ATG GTG CG
<i>TdTomato</i>	AAG GGA GCT GCA GTG GAG TA	CCG AAA ATC TGT GGG AAG TC
<i>hSS2</i>	GTT ATC AGT AAG GGA GCT GCA GTG G	AAG ACC GCG AAG AGT TTG TCC TC AND GGC GGA TCA CAA GCA ATA ATA ACC
<i>Ctnnb1 Ex3Fl</i>	GGT AGT GGT CCC TGC CCT TGA CAC	CTA AGC TTG GCT GGA CGT AAA CTC
<i>Ctnnb1 WT</i>	GGT AGG TGA AGC TCA GCG CAG AGC	ACG TGT GGC AAG TTC CGC GTC ATC C