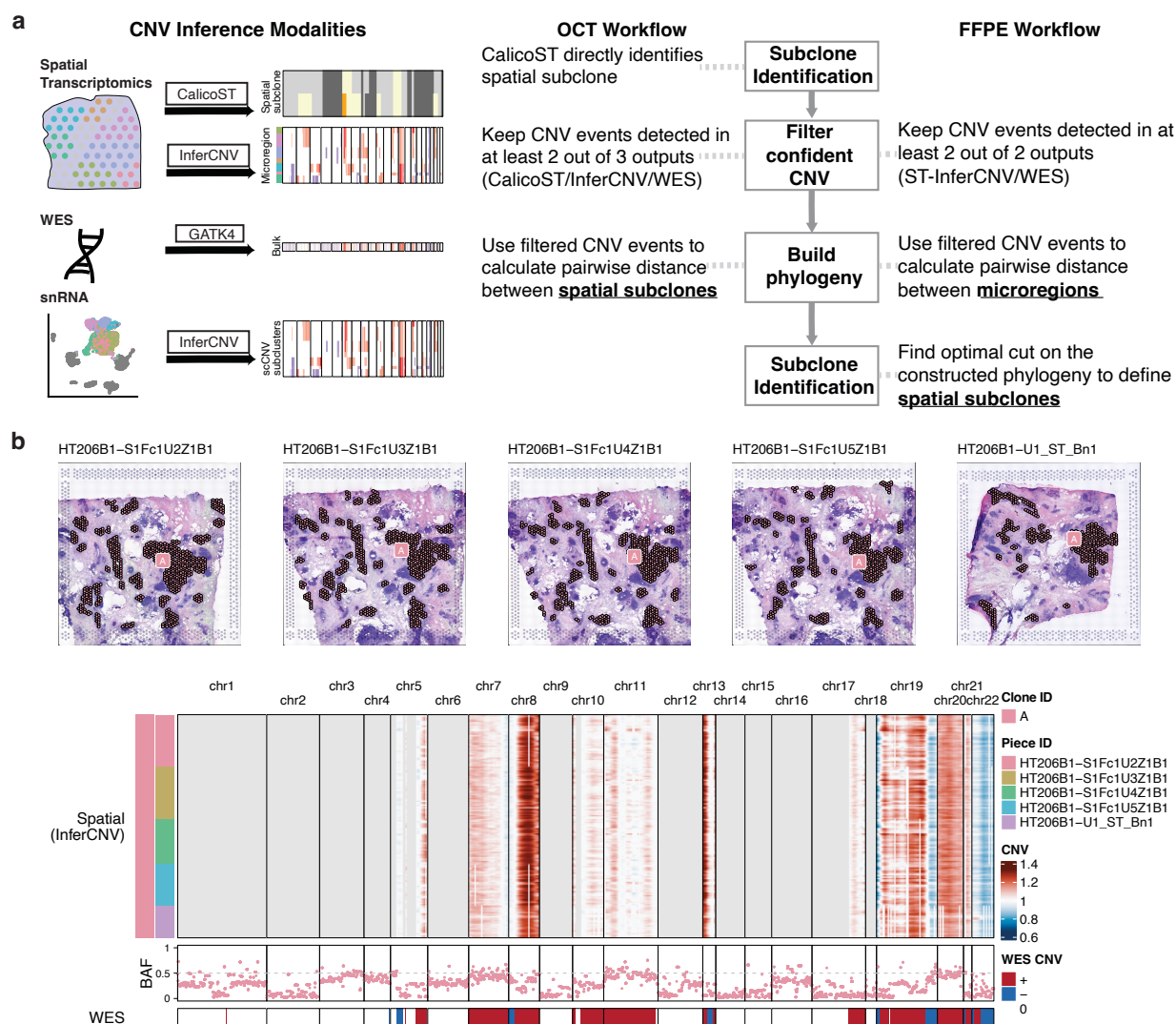

Supplementary information

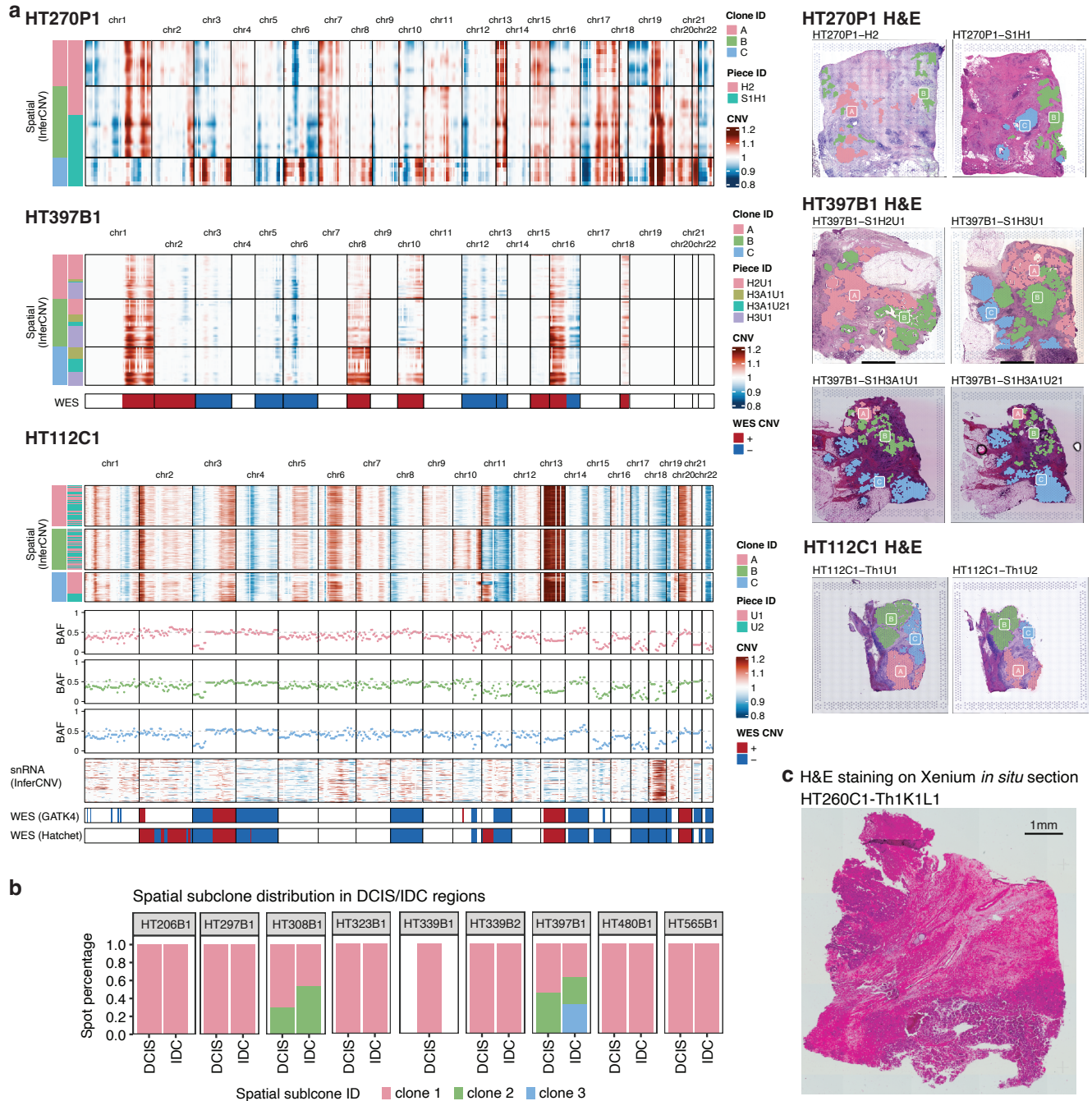
Tumour evolution and microenvironment interactions in 2D and 3D space

In the format provided by the
authors and unedited



Supplementary Figure 1. Spatial subclone identification workflow

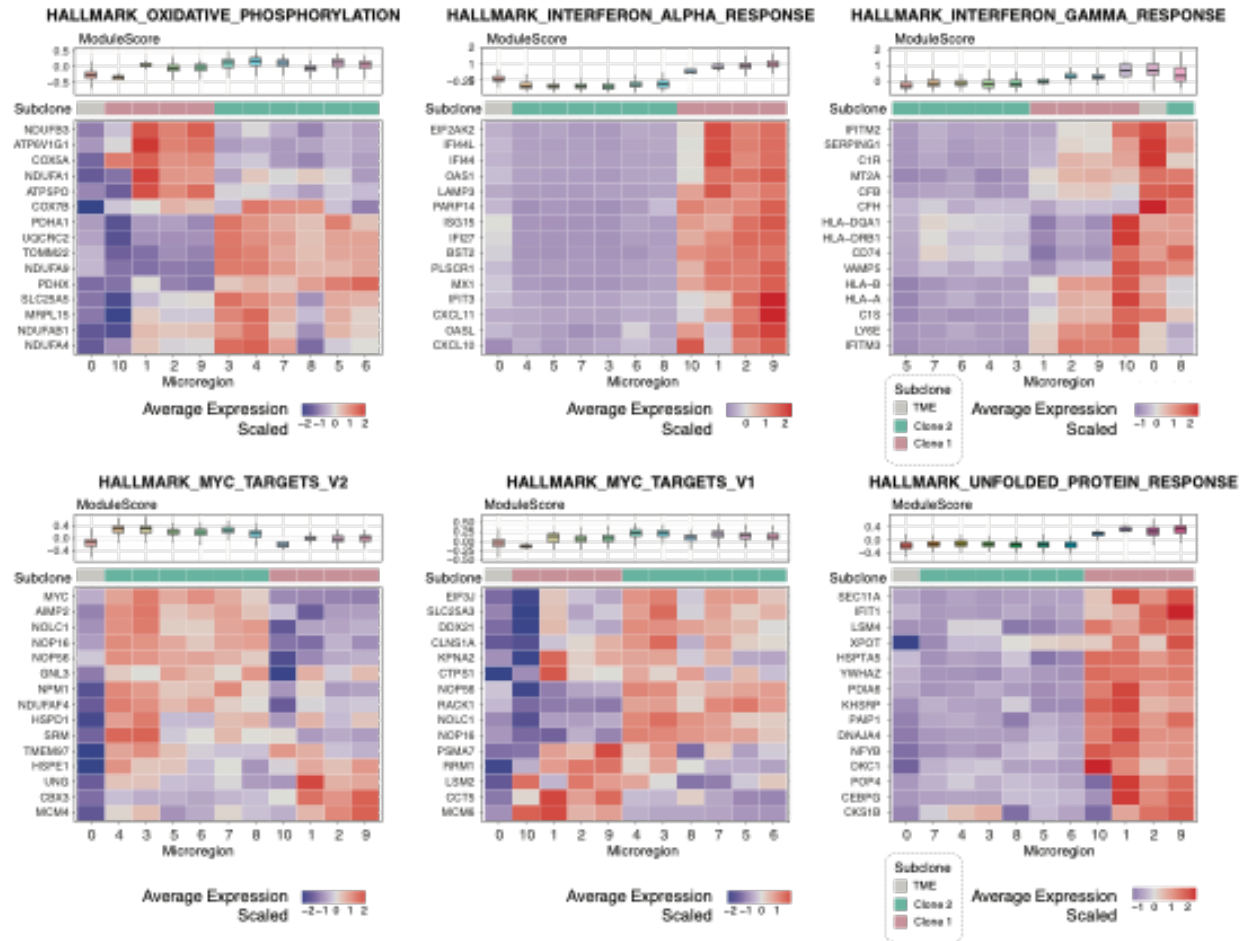
a, Spatial subclone identification workflow. Samples with variant information (Optimal Cutting Temperature) and samples with only transcript count information (Formalin-Fixed Paraffin-Embedded) samples have different orders of processing to achieve the highest sensitivity. **b**, HT206B1, a primary BRCA sample, showing a single clone across 5 sections.



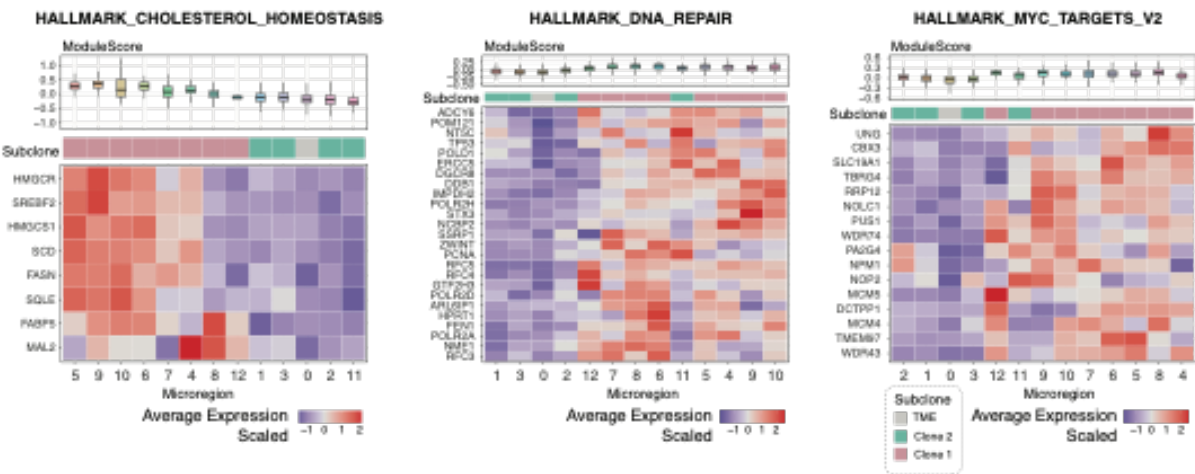
Supplementary Figure 2. Additional examples for spatial subclones

a, Full genome-wide CNV profile demonstrating subclonal structures, related to Fig. 2. **b**, Spatial-subclone distribution in breast cancer samples between DCIS and IDC morphologies. Bar height represents the fraction of each morphology (DCIS or IDC) belonging to each spatial subclone. **c**, Hematoxylin and eosin (H&E) staining of HT260C1 Xenium section.

a HT268B1-Th1H3U1 (Breast Metastasis)

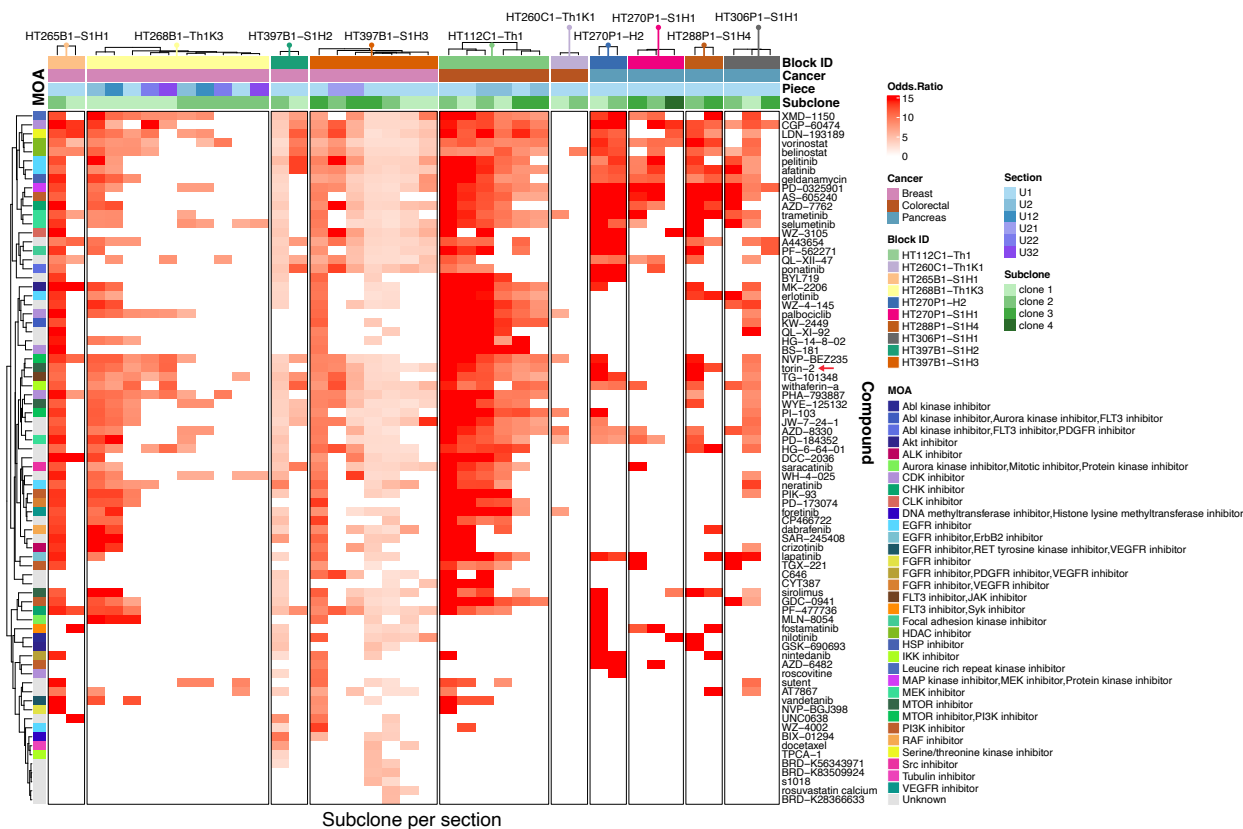


b HT260C1-Th1K1U1 (Colorectal Metastasis)



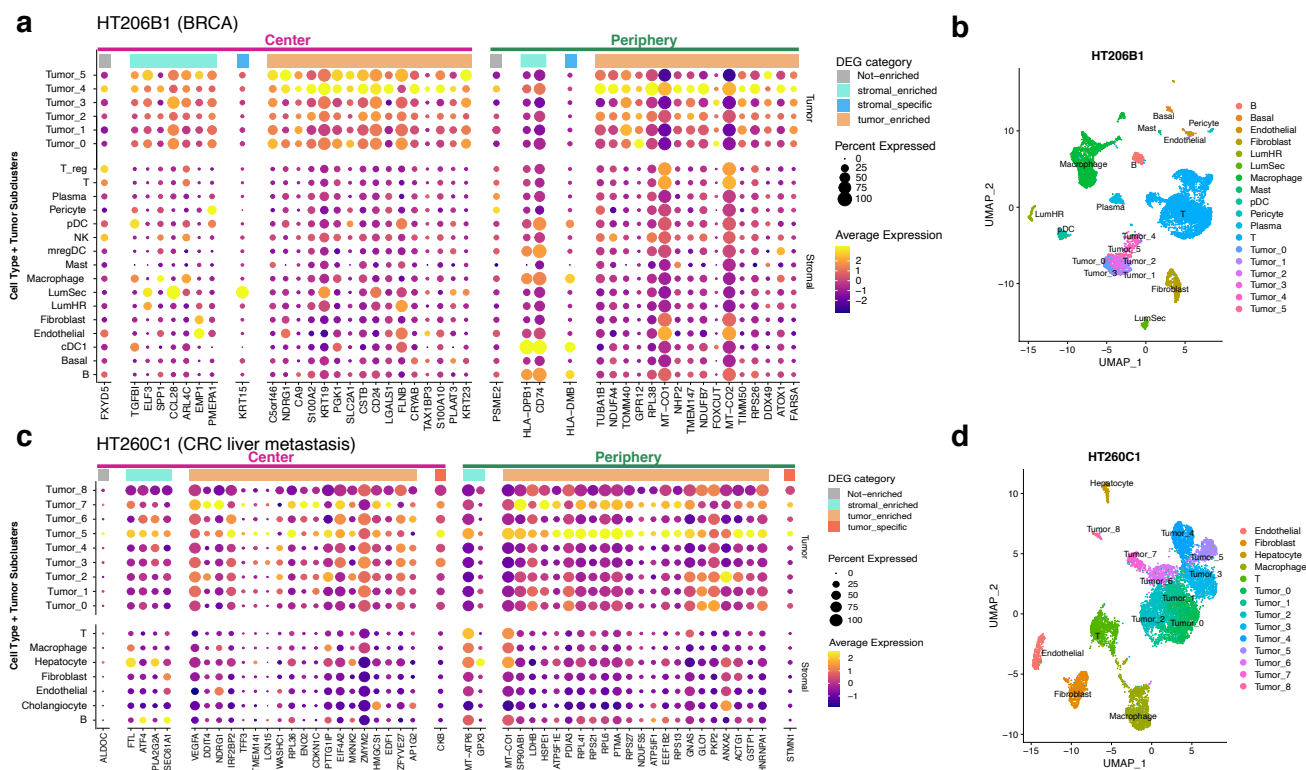
Supplementary Figure 3. Transcriptional disparities in tumor microregions

Average expression heatmap of genes in selected hallmark pathways with distinct subclone-specific expression pattern in **a**, HT268B1-Th1H3U1, and **b**, HT260C1-Th1K1U1. The box plot's center line represents the median, with the lower and upper hinges indicating the first and third quartiles. Whiskers extend to the highest and lowest values within 1.5 times the interquartile range (IQR) from the hinges.



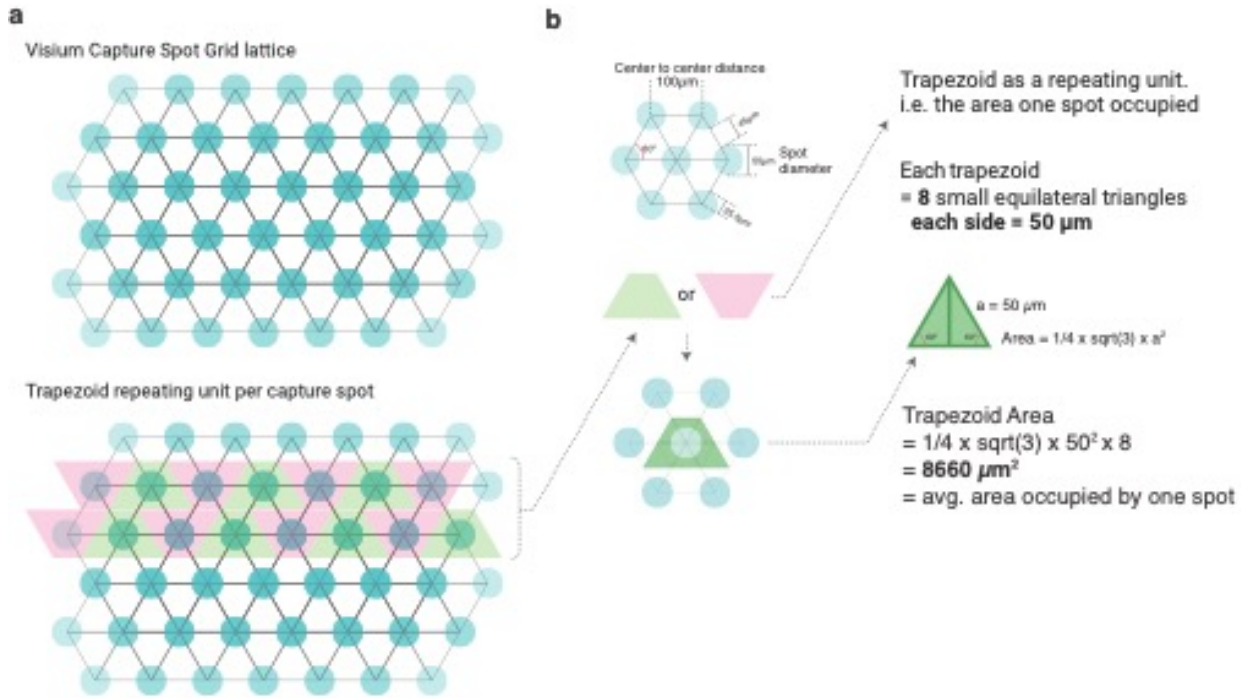
Supplementary Figure 4. Spatial subclone-specific treatment response analysis

Enrichment of subclone DEGs in compound perturbation dataset LINCS L1000 downregulated genesets. Heatmap values are “Odds.Ratio” which indicates geneset overlaps. Each row is one compound annotated by the mechanism of action (MOA). Column denotes subclones from each section.



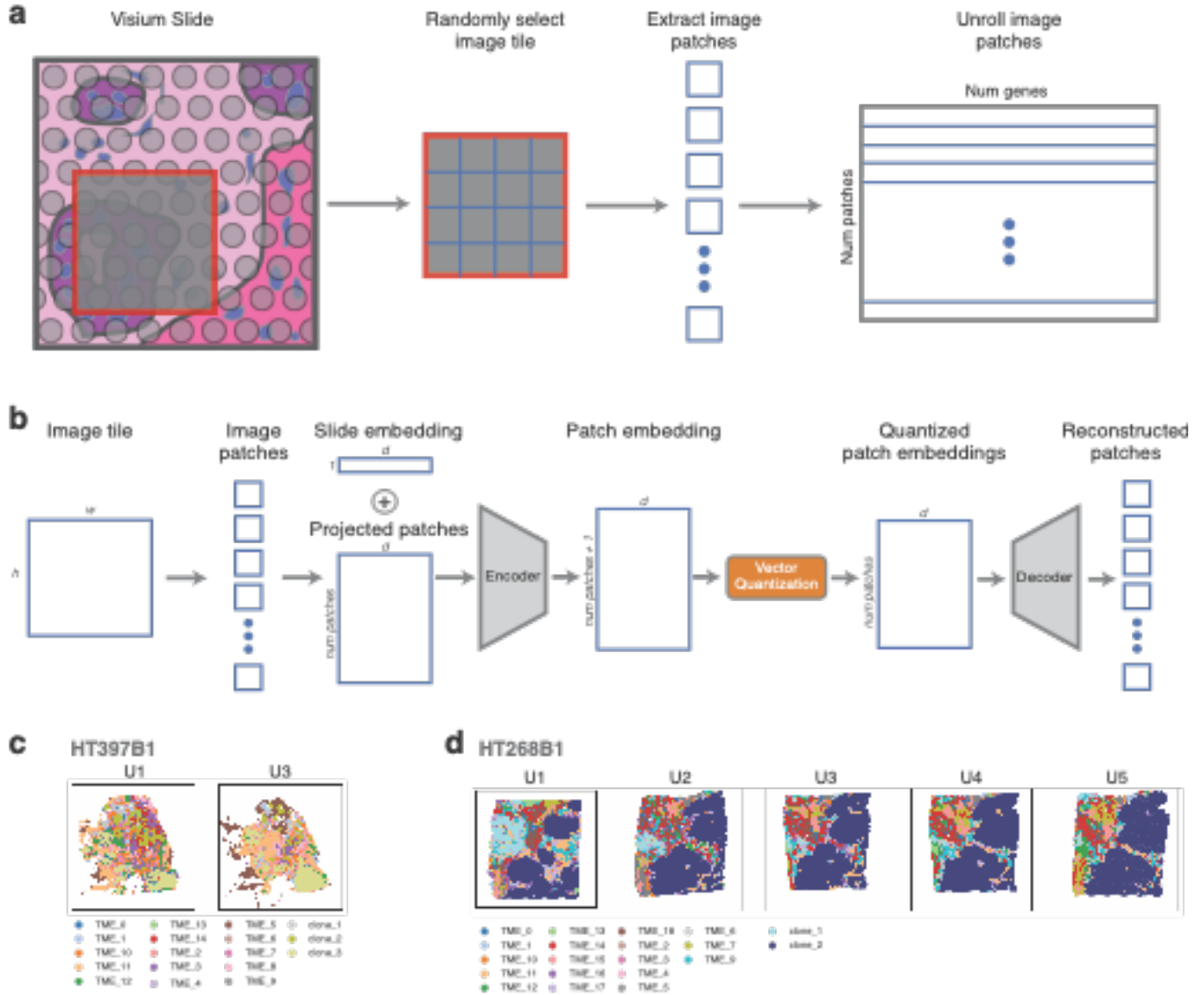
Supplementary Figure 5. Differential biological processes between the tumor core and leading edge

a, Dotplot of categorized DEG in the colorectal liver metastasis sample HT260C1. **b**, UMAP of tumor subcluster and non-tumor cell types in matching snRNA-seq. **c**, Dotplot of categorized DEG in the primary breast cancer sample HT206B1. **d**, UMAP of tumor subcluster and non-tumor cell types in matching snRNA-seq.



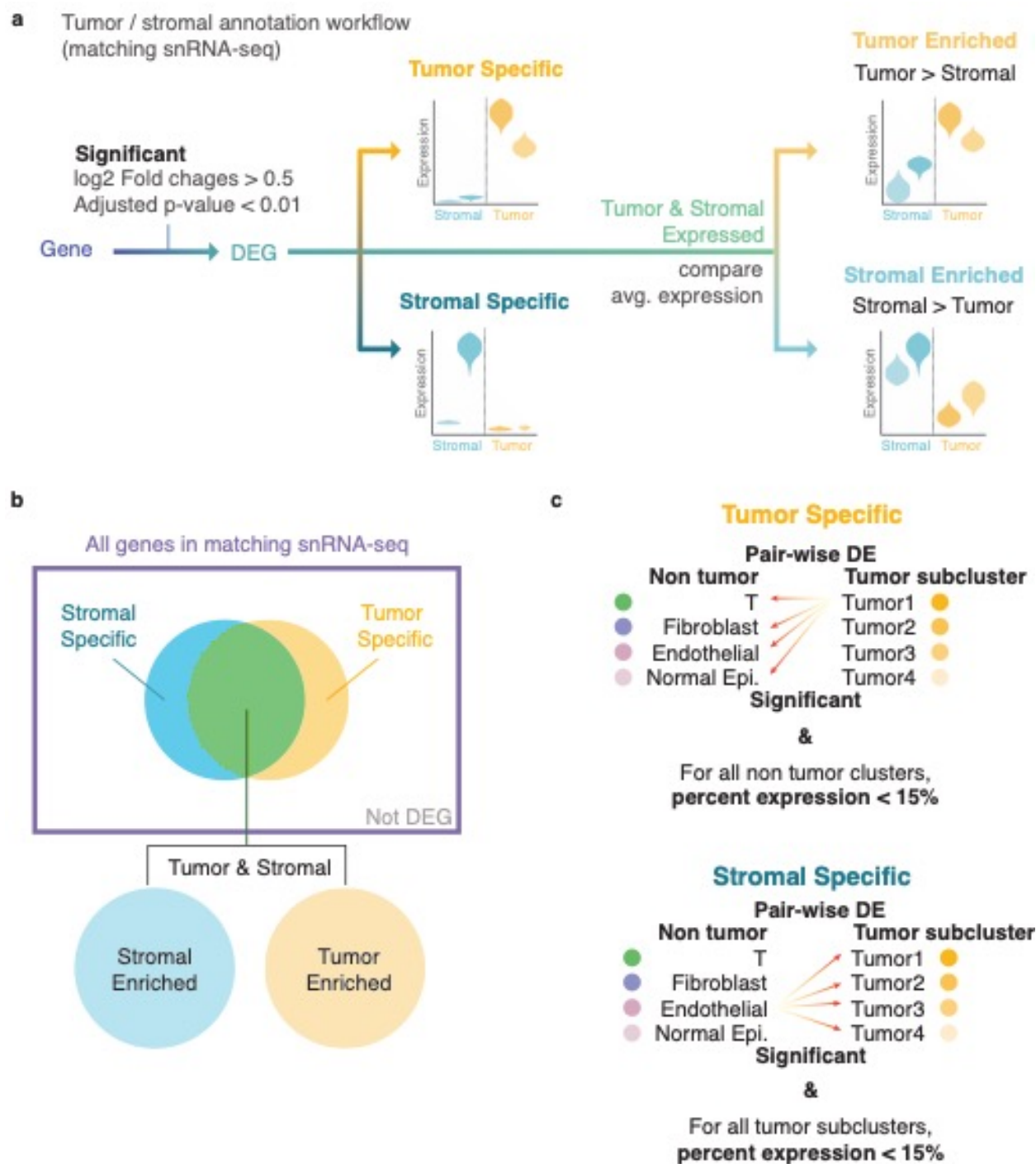
Supplementary Figure 6. Average spot area and microregion size calculation

a, the spatial orientation of Visium ST spots in lattice format, and the repeating trapezoid unit. **b**, (Left) size, center-to-center, and space between each spot. (Right) Repeating trapezoid unit calculation to deduce the average area of one spot.



Supplementary Figure 7. Neighborhood identification Model architecture

a, Schema of tile sampling workflow to create image patches that are input to the ViT autoencoder. **b**, an overview of the ViT autoencoder responsible for identifying patch neighborhoods. First, the image tile is split into patches. Each patch is then projected to the transformer embedding dimension and combined with a slide token representing the section the image tile is sampled from. The patches are then encoded with a variational autoencoder, after which the latent patch embeddings are quantized. After quantization, each token is then used to reconstruct the patches' input expression profile. Neighborhoods and subclone annotations for ST sections of **c**, HT397B1 (U1 and U3) and **d**, HT268B1.



Supplementary Figure 8. Tumor intrinsic and non-tumor gene categorization

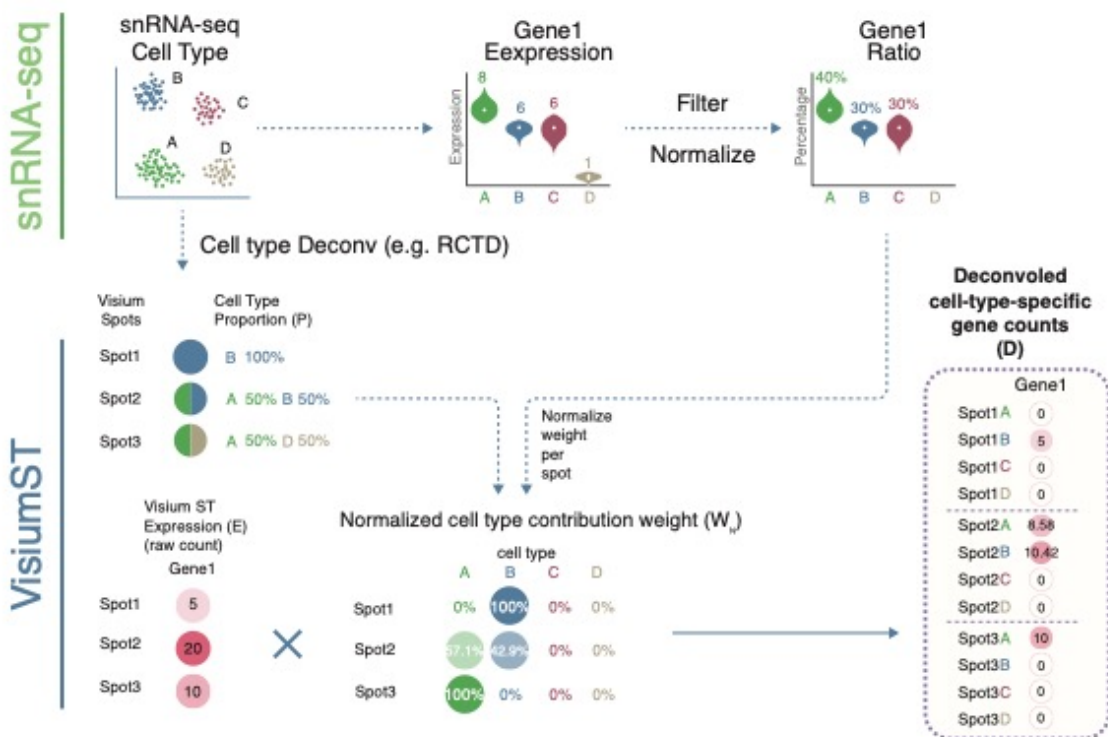
a, Workflow diagram of filtering and categorizing DEG as tumor-specific, stromal-specific, tumor-enriched, and stromal-enriched categories. DEGs that are not tumor or stromal-specific are tumor/stromal enriched. An enriched

gene is categorized as tumor-enriched if the highest average expression of that gene in the tumor is higher than in the non-tumor cluster. Vice versa for those categorized as stromal enriched genes. **b**, Venn diagram of relationships among DEG categories. **c**, diagram of pair-wise differential expression analysis between stromal and tumor clusters. A gene is considered tumor/stromal specific only if 1. It's a pair-wise DEG in at least one tumor subcluster compared to all non-tumor cell types, and 2. No nontumor cell type has a percent expression of the gene greater than 15%. Similarly, a stromal-specific gene must be pair-wise DEG to all tumor subclusters and all tumor subclusters have less than 15% gene expression.

a

Deconvolute Visium ST spot cell-type-specific expression based on

- snRNA-seq expression & Cell type deconvolution proportion

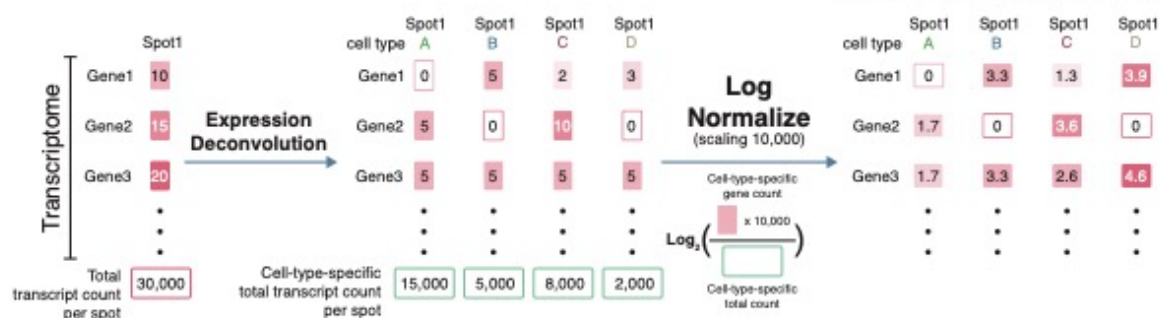


b

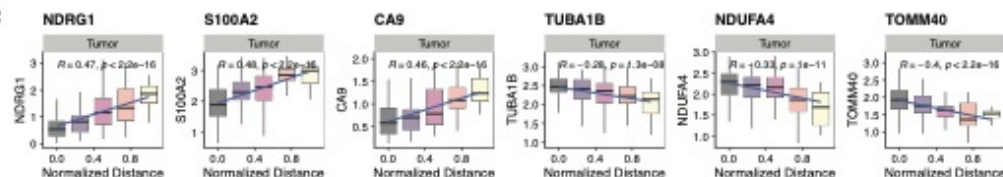
Transcript count

Transcript count

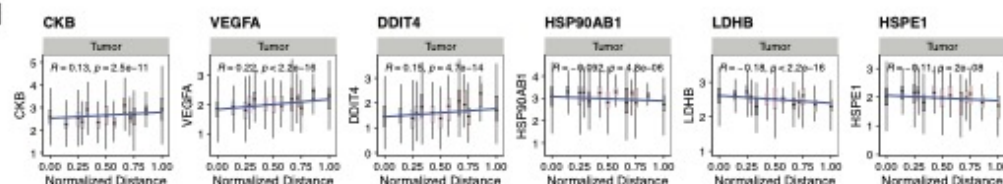
Normalized expression
(to cell-type-specific library size)



C

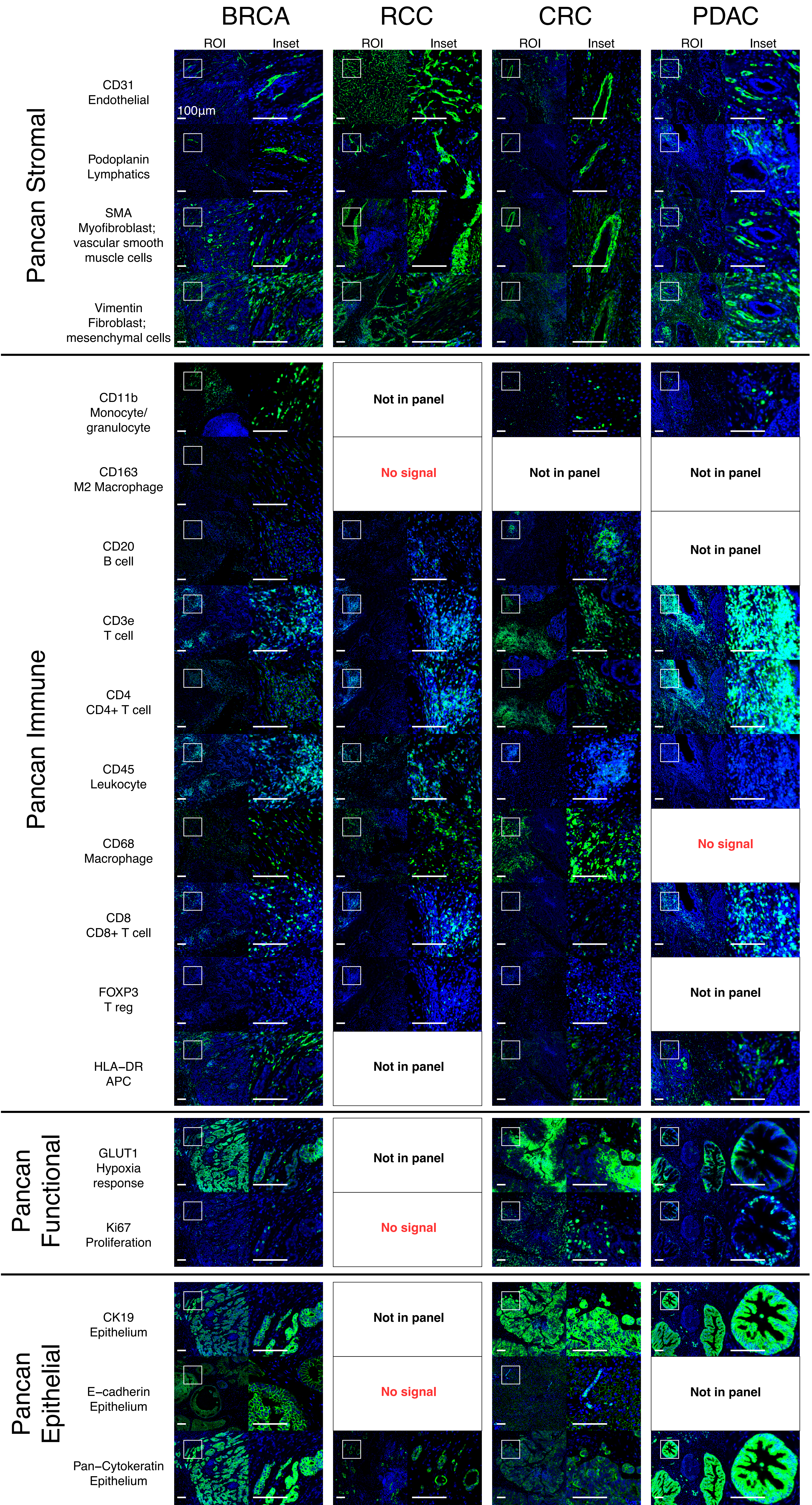


d

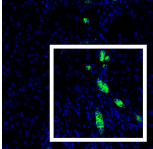
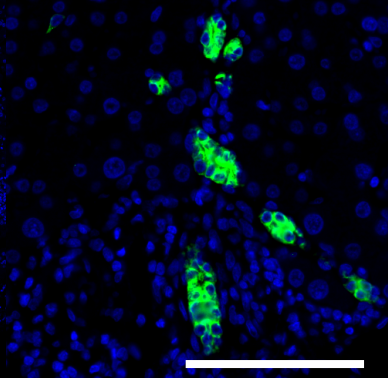
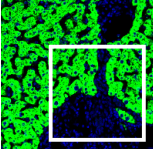
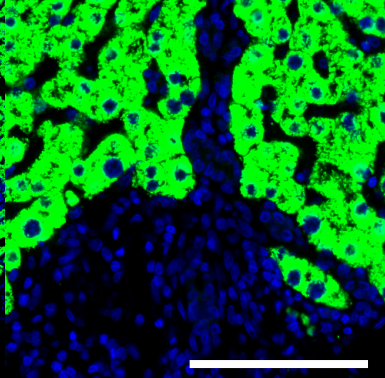
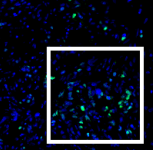
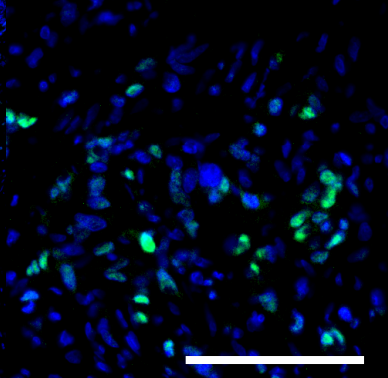
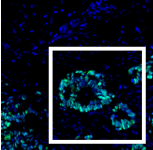
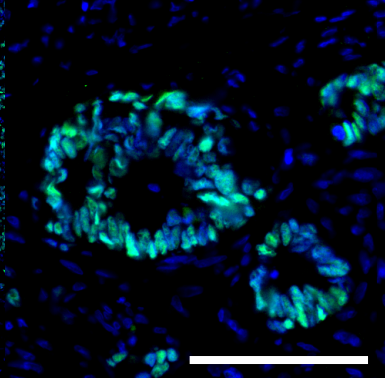
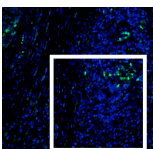
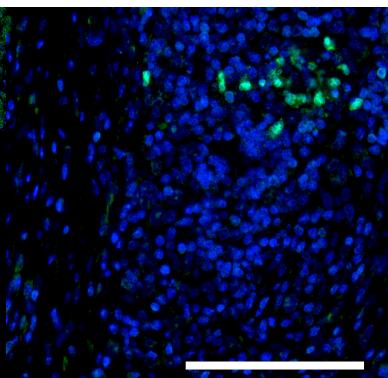
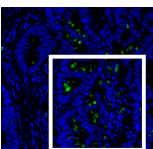
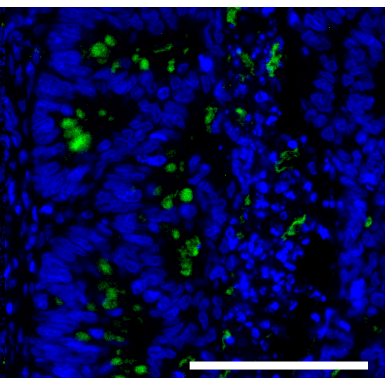
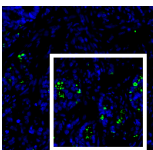
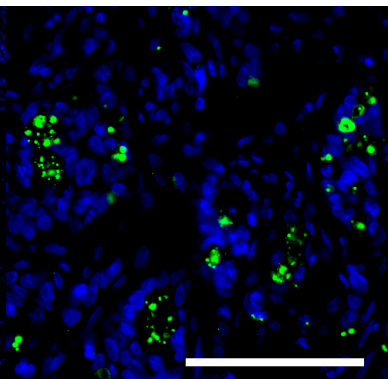
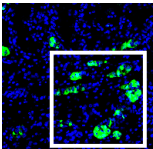
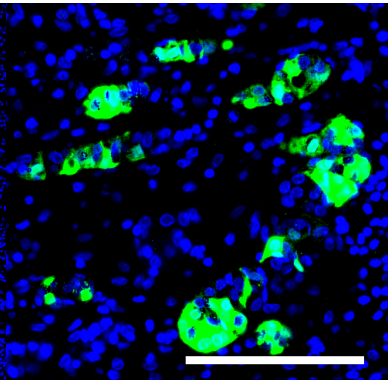
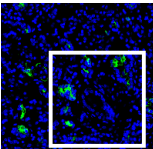
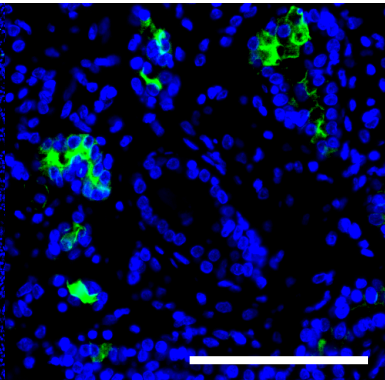
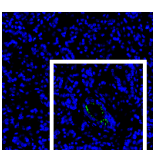
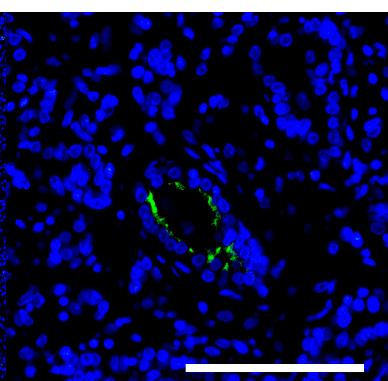
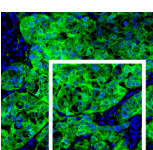
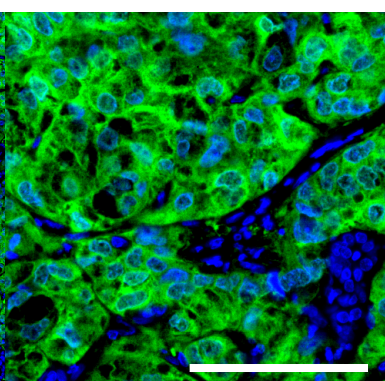
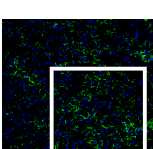
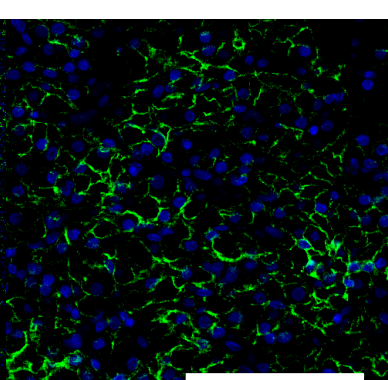
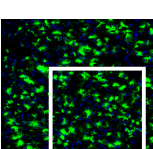
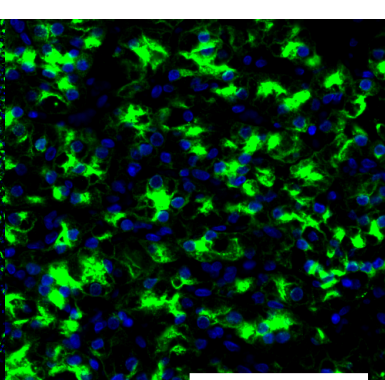
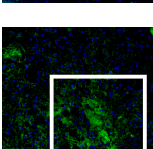
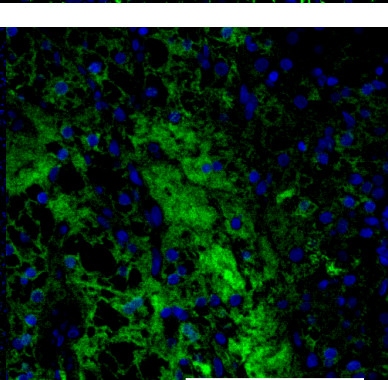
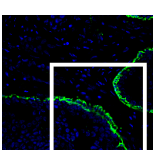
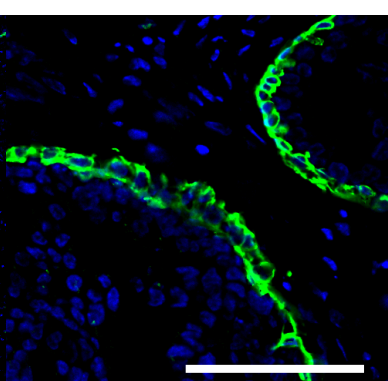
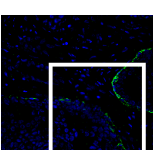
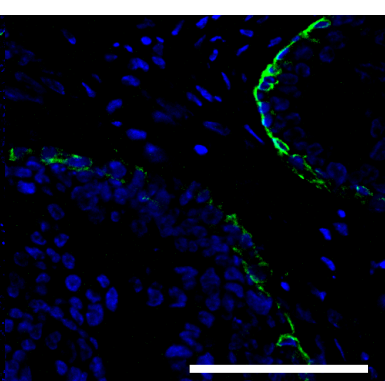
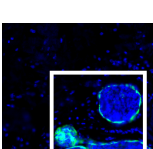
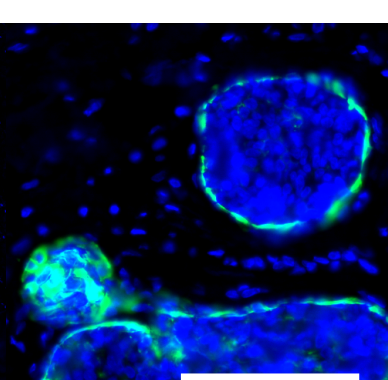
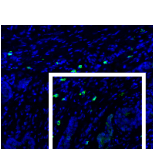
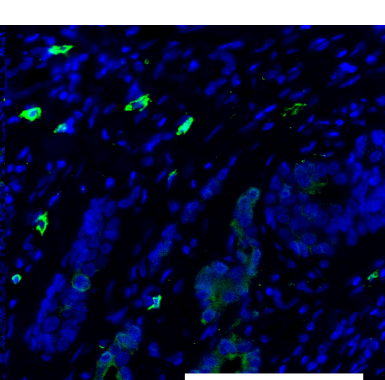
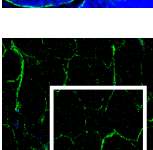
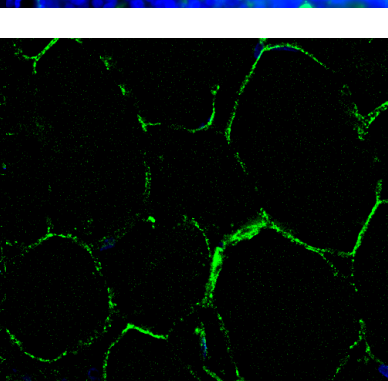
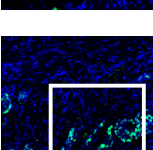
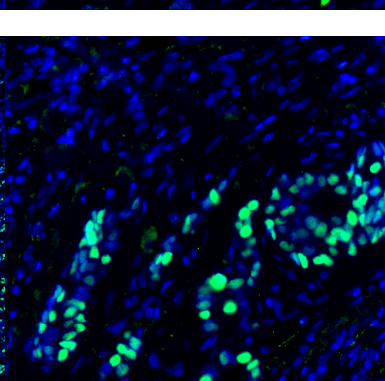
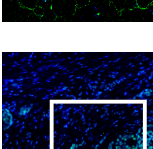
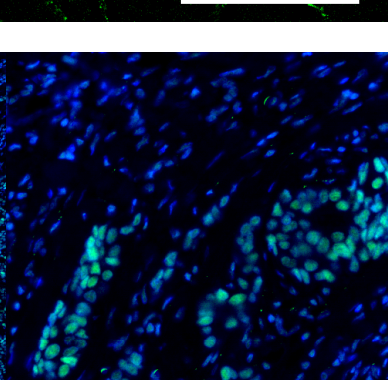
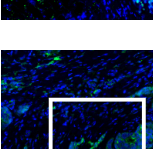
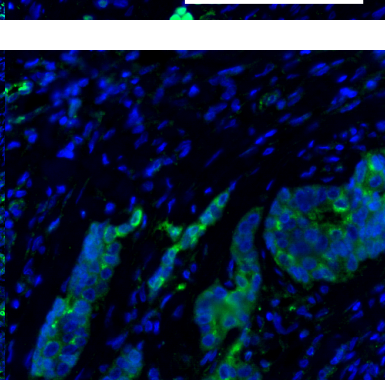
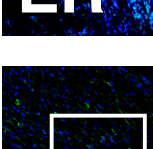
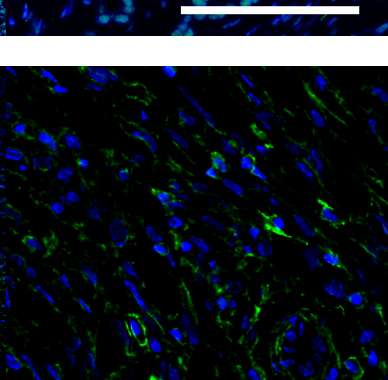
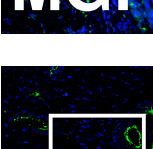
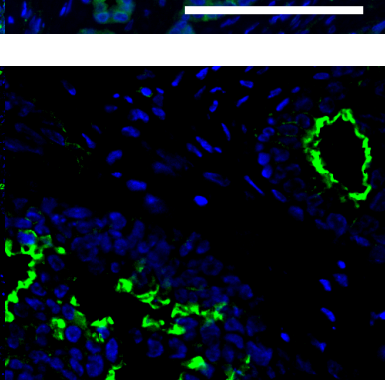
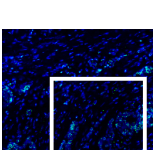
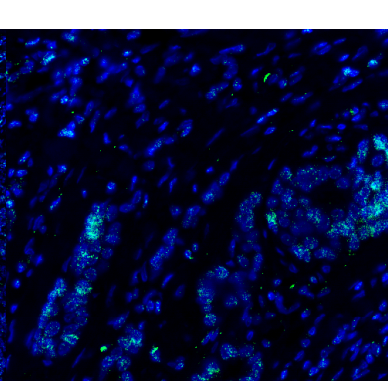
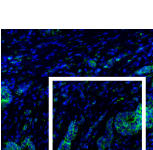
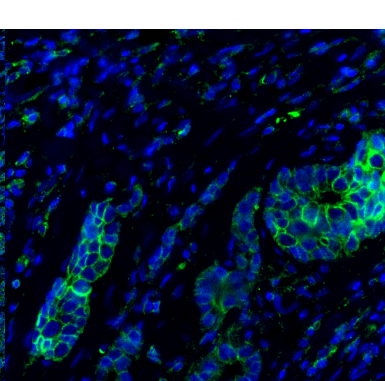
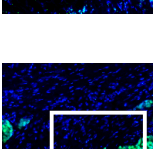
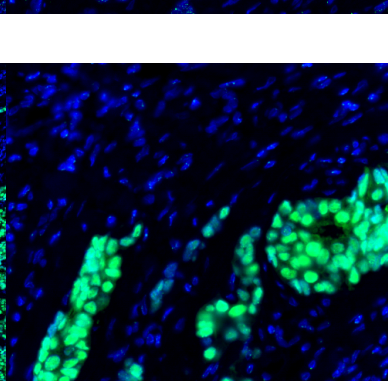
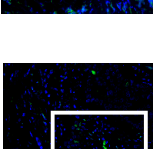
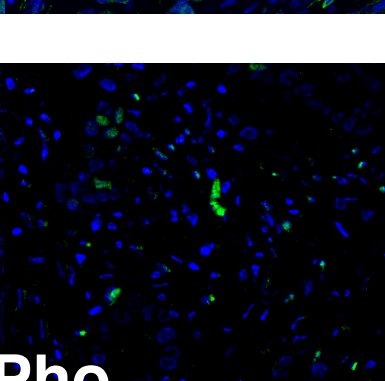
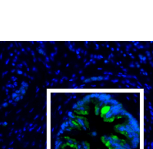
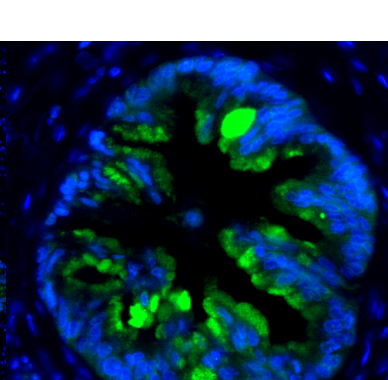


Supplementary Figure 9. Spatial Expression Deconvolution

a, The diagram of expression deconvolution workflow. This workflow utilizes spot-level cell type deconvolution (using tools such as RCTD), and average expression per cell type from matching snRNA-seq data to deconvolve the expression per Visium spot. In short, for a given Gene1 in Spot1, the cell type proportion (P) multiplied by the cell type-specific expression level in snRNA-seq data (Q) gives the weight (W) that determines the contribution of a given cell type to that Gene1 in Spot1. The raw weights from all cell types in Spot1 then are normalized to 1 to form normalized weight (WN). This weight is then multiplied by the spot level expression from ST (E) to get the final sub-spot level expression per cell type (D). **b**, Cell-type-specific expression normalization. Layer expression trend plots in **c**, HT206B1, and **d**, HT260C1 used deconvoluted expression. The box plot's center line represents the median, with the lower and upper hinges indicating the first and third quartiles. Whiskers extend to the highest and lowest values within 1.5 times the interquartile range (IQR) from the hinges.



Supplementary Figure 10. CODEX pan-cancer antibody panel with representative single-channel views (combined with DAPI in blue). Antibodies not included in a specific panel are labeled “not in the panel”, while antibodies showing no signal in a specific cancer type are labeled “No signal”. Scale bars are 100µm.

CRC specific	CK7 Cholangiocyte (Liver compartment)	ROI  CK7 100μm Inset 	Hep-Par-1 Hepatocyte (Liver compartment)	ROI  Hep-Par-1 Inset 
	P21 (CRC functional)	 P21 Inset 	SOX9 (CRC functional)	 SOX9 Inset 
	CCL2 (CRC functional)	 CCL2 Inset 	MUC2 (CRC functional)	 MUC2 Inset 
	PAI1 (CRC functional)	 PAI1 Inset 		
RCC specific	CALB1 Collecting duct; distal tubules (Kidney compartment)	 CALB1 Inset 	AQP1 Proximal tubule (Kidney compartment)	 AQP1 Inset 
	AQP2 Collecting duct (Kidney compartment)	 AQP2 Inset 	UCHL1 Tumor marker (ccRCC functional)	 UCHL1 Inset 
	CA9 Tumor marker (ccRCC functional)	 CA9 Inset 	LRP2 Proximal tubule; Tumor marker (ccRCC functional)	 LRP2 Inset 
	CP Tumor marker (ccRCC functional)	 CP Inset 		
BRCA specific	CK14 Breast myoepithelial (Breast compartment)	 CK14 Inset 	CK17 Breast myoepithelial (Breast compartment)	 CK17 Inset 
	CK5 Breast myoepithelial (Breast compartment)	 CK5 Inset 	cKit Luminal progenitor (Breast compartment)	 cKit Inset 
	CD36 Adipocyte (Breast compartment)	 CD36 Inset 	PR Subtype (BRCA functional)	 PR Inset 
	ER Subtype (BRCA functional)	 ER Inset 	MGP (BRCA functional)	 MGP Inset 
	CD44 (BRCA functional)	 CD44 Inset 	MLPH (BRCA functional)	 MLPH Inset 
	COX6c (BRCA functional)	 COX6c Inset 	PLAT/tPA (BRCA functional)	 PLAT Inset 
	GATA3 (BRCA functional)	 GATA3 Inset 	p-Histone H3 (BRCA functional)	 Histone H3 Pho Inset 
PDAC specific	TFF1 (PDAC functional)	 TFF1 Inset 		

Supplementary Figure 11. CODEX cancer-specific antibody panel with representative single-channel views (combined with DAPI in blue). Antibody name, expected staining pattern, and the relevant cancer type are indicated to the left of each image. Scale bars are 100μm.