**Figure 1. Overview of spatial molecular imaging (SMI) experiment utilizing HCC FFPE sections.** (**A**) Experimental workflow of spatial transcriptomic profiling experiment. (**B**) Serial section H&E staining (left) and corresponding original immunofluorescence staining (right) for core tissue samples used in the SMI experiment from TMA2. For TMA1, see Figure S3. White rectangles depict fields of view (FOVs), within which SMI was conducted. Green: pan-cytokeratin, red: CD45, yellow: CK8/18, cyan: CD298/B2M. (**C**) Immunofluorescent staining used for cell segmentation in a representative FOV: TMA2, FOV60. (**D**) Cellpose cell segmentation results (left) and representation of SMI transcript detection (right) used to identify tumor and immune cells in the same FOV as in 1C.

**Figure 2. Subcellular spatial molecular imaging (SMI) accurately identifies cell types in HCC tumor microenvironment.** (**A**) Schematic of cell-typing workflow, which included both manual annotation and supervised clustering, applied to TMA1. The same workflow was applied to TMA2 (see methods). (**B**) Final cell type annotations visualized on the UMAP embedding for the complete SMI dataset (left). Cells mapped back onto their physical locations in space, pseudo-colored by final cell type annotation for the same FOV in 1C (right). (**C**) Dot plot depicting expression of marker genes found algorithmically (see methods) for each of the identified cell type clusters. (**D**) Bar plot depicting the composition of each FOV in the complete SMI dataset in terms of major cell type categories.

**Figure 3. HCC expression of LINE1 is associated with a de-differentiated, stem-like state.** (**A**) Bar plot depicting the definition of LINE1-ORF1 high and low patient groupings in the SMI dataset. The dashed line represents the upper tercile for net tumor cell LINE1-ORF1 counts per million (CPM) across all patients. (**B**) Volcano plot summarizing differential expression analysis comparing tumor cells in LINE1-ORF1 high patients and LINE1-ORF1 low patients, as defined in 3A. P-values are FDR-adjusted. Significance threshold = 0.05. (**C**) Heatmap depicting average normalized expression of selected differentially expressed genes identified in 3B among tumor cells in each patient. Values have been row scaled for visualization. (**D**) Box plots depicting the definition of LINE1-ORF1 high and low tumor cell groupings within each patient in the SMI dataset. The black crossbars represent the upper tercile for normalized LINE1-ORF1 expression within each patient. (**E**) Volcano plot summarizing differential expression analysis comparing LINE1-ORF1 high tumor cells and LINE1-ORF1 low tumor cells in a patient-paired manner. P-values are FDR-adjusted. Significance threshold = 0.05. LINE1-ORF1 not shown. (**F**) Patient-paired differences in total normalized expression of selected differentially expressed genes identified in 3E between LINE1-ORF1 high and low tumor cells. The crossbars represent means. (**G**) Gene set enrichment analysis plots generated from the results depicted in 3E.

**Figure 4. High LINE1 expression within HCC tissue leads to a disorganized, sparse immune microenvironment.** (**A**) Patient-paired differences in LINE1-ORF1 high and low tumor cells’ colocalization patterns with themselves and one another. ER = enrichment ratio (see methods). P-values computed by paired Wilcoxon Rank Sum Test and FDR-adjusted. For other cell types, see Figure S6a. (**B**) Representative spatial maps for two FOVs depicting tumor niches comprised mainly of cells with the same LINE1-ORF1 group identity, as predicted by 4A. (**C**) Neighborhood enrichment analysis depicting pairwise colocalization for all cell type pairs in LINE1-ORF1 high patients (left) and LINE1-ORF1 low patients (right). (**D**) Randomization testing for difference in immune organization score between LINE1-ORF1 high and low patients. The red dashed line represents the observed difference which was compared to the shown empirical null distribution (see methods) to compute the reported p-value. (**E**) Representative immunofluorescent staining images and corresponding spatial maps for four FOVs illustrating differences in immune cell organization in LINE1-ORF1 high patients (left) and LINE1-ORF1 low patients (right).

**Figure 5. Validation of LINE1 expression by RNA-ISH.** (**A**) Representative RNA-ISH image showing LINE1 expression. (**B**) Jitter plot depicting the definition of LINE1 high and low patient groupings in the RNA-ISH dataset. The dashed line represents the upper tercile for LINE1 counts per µm2 across all patients. The crossbars represent means. P-value computed by Wilcoxon Rank Sum Test. (**C**) Association of LINE1 RNA-ISH quantification with HERVH, HERVK, and HSATII RNA-ISH quantification. P-values computed by Wald test. (**D**) Kaplan-Meier overall survival analysis for LINE1 high patients versus LINE1 low patients (p = 0.01 by log-rank test).

**Joseph’s drafts:**

**Figure 1. Overview of spatial molecular imaging (SMI) experiment utilizing HCC FFPE sections.** (**A**) Experimental workflow of spatial transcriptomic profiling experiment. (**B**) Schematic showing subset of TMA core tissue samples utilized in one spatial molecular imaging (SMI) “run.” H&E staining of a consecutive section on the left and immunofluorescent staining (green: pan-cytokeratin, red: CD45, yellow: CD68, cyan: beta-2-microglobulin) of the imaged slide section on the right. (**C**) Immunofluorescent “morphology marker” staining on a representative field of view (FOV). (**D**) Cell boundary demarcation and tumor versus immune cells shown in the same FOV as shown in C.

**Figure 2. Subcellular spatial molecular imaging (SMI) accurately** **identifies cellular subtypes in HCC tumor microenvironment**. (**A**) Schema of combined supervised and unsupervised computational workflow for cell type identification. (**B**) UMAP plot of cell types – with embedded spatial coordinate metadata allowing for mapping back onto the 2D slide surface – identified in the HCC SMI dataset. (**C**) Bubble plot showing expression of marker genes for each of the identified cell types. (**D**) Proportion of each major cell type category present within each FOV of the dataset.

**Figure 3. HCC expression of LINE-1 is associated with a de-differentiated, stem-like state**. (**A**) Bar plot of scaled LINE1-ORF1 counts within HCC cancer cells in “high” (top tercile) and “low” (middle plus bottom tercile) groups. (**B**) Volcano plot comparing LINE1-high and LINE1-lowcore (entire sample) groups. Statistical threshold set at false discovery rate corrected p-value of <0.05. (**C**) Heat map showing scaled expression of differentially expressed genes between LINE1-high and LINE1-low cores. (**D**) Box and whisker plots for each core showing LINE1-high and LINE1-low cores showing single cell expression of LINE-1, subdivided in each core into high and low expressing groups. (**E**) Volcano plot comparing LINE1-high and LINE1-low single cells. Statistical threshold set at false discovery rate corrected p-value of <0.05. (**F**) Normalized expression of HSATII, HERV-K, POU5F1 (Oct4), LEFTY1, and SERPINA3 between LINE1-high and LINE1-low single HCC cells. (**G**) Gene set over-representation analysis of LINE1-high versus LINE1-low single HCC cells.

**Figure 4. High LINE-1 expression within HCC tissue leads to a disorganized, sparse immune microenvironment**. (**A**) Paired plots of niche enrichment ratios of LINE-high and LINE1-low cells with one another. (**B**) Representative images of HCC LINE1-high (red dots) and LINE1-low (blue dots) neighborhoods, with all other cell types represented as white dots. (**C**)Heatmaps of enrichment ratios of the microenvironment niches of LINE1-high tumor tissues (left) versus LINE1-low tumor tissues (right). (**D**) Observed immune organization structure (red dashed vertical line) versus simulated random distrubtion of \_\_\_\_\_\_. (**E**) Representative images of LINE1-high (left two columns) versus LINE1-low (right two columns), with immune cells represented as red dots and all other cells as white dots (first row), with paired immunofluorescent images shown in the second row (red: CD45, yellow: keratin 8/18, blue: DAPI).

**Figure 5. Validation of LINE-1 expression by RNA-ISH**. (**A**) Representative LINE-1 RNA-ISH image. (**B**) Plot of LINE1 RNA ISH quantitative expression in the “high” (upper tercile) group versus the “low” (bottom terciles) group. (**C**) Correlation of HERV-H and HERV-K with LINE-1 expression by RNA-ISH. (**D**) Kaplan-Meier overall survival analysis for patients with LINE1-high (red) versus LINE1-low (blue) tumor tissues (p = 0.01 by log-rank test).