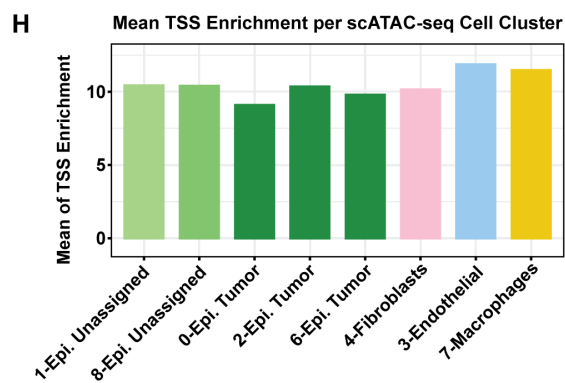
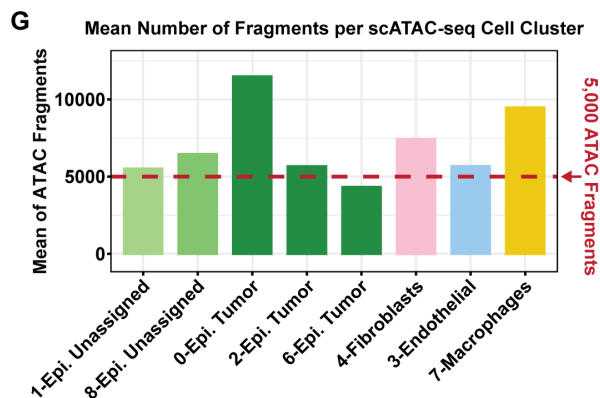
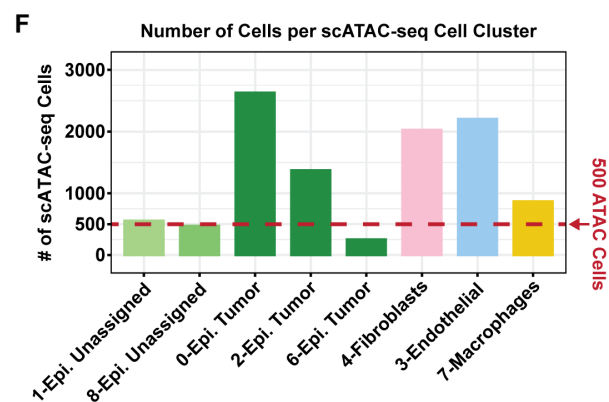
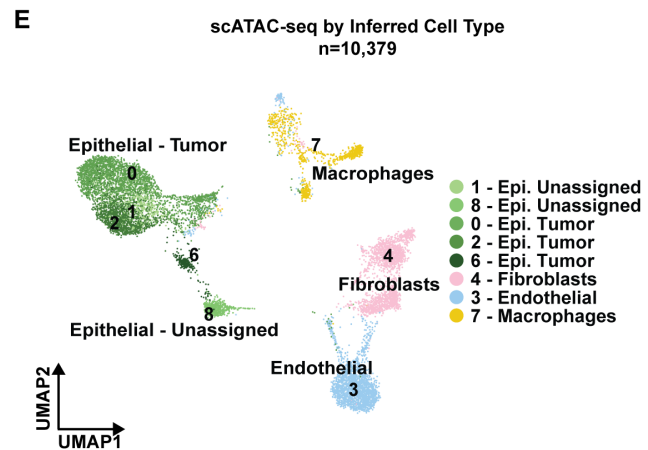
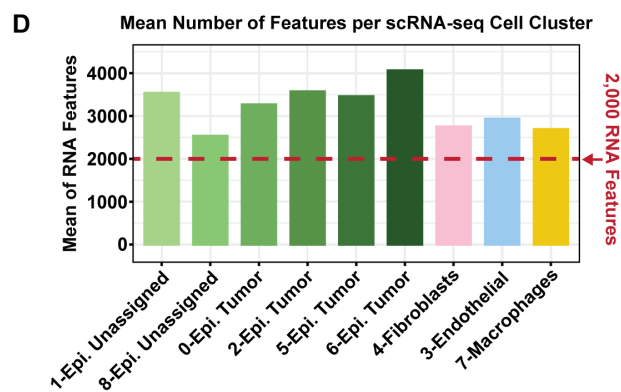
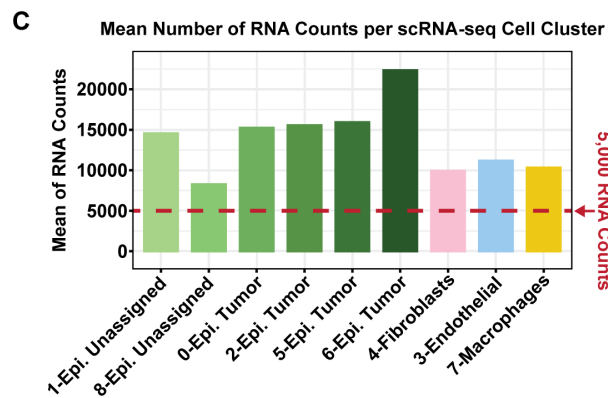
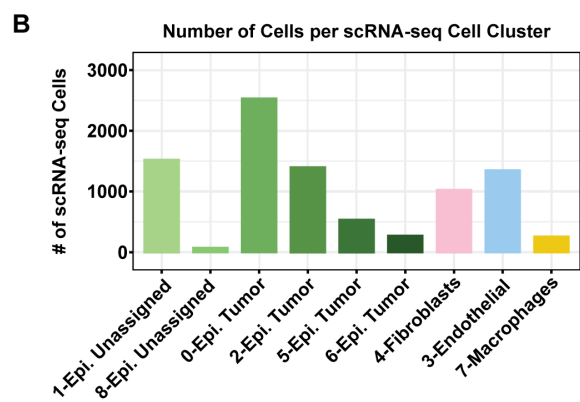
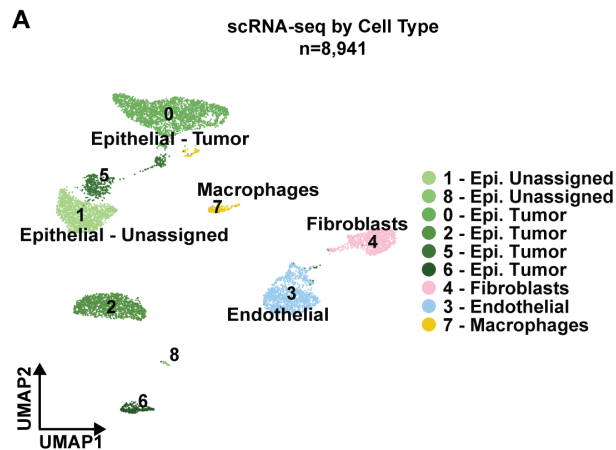


Single-Cell Transcriptional and Epigenetic Profiles of Male Breast Cancer Nominate Salient Cancer-Specific Enhancers

Hyunsoo Kim^{1,6}, Kamila Wisniewska^{1,6}, Matthew J. Regner^{1,2}, Aatish Thennavan^{1,3}, Philip M. Spanheimer^{1,4} and Hector L. Franco^{1,2,5*}

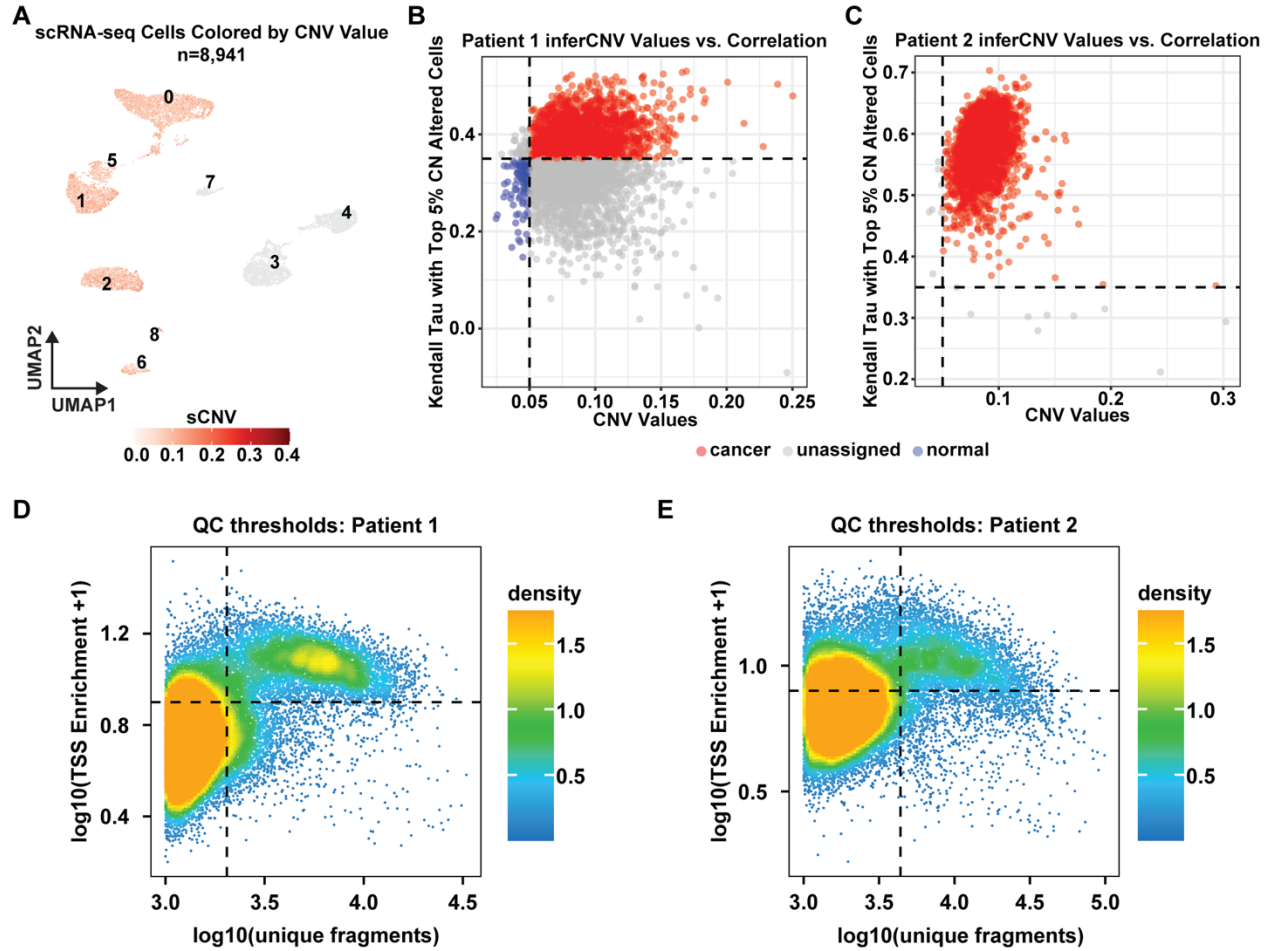
Supplementary Information

1. Supplementary Figure S1. Quality metrics for scRNA-seq and scATAC-seq.
2. Supplementary Figure S2. CNV Thresholds for scRNA-seq Cells and QC Thresholds for scATAC-seq Cells
3. Supplementary Figure S3. Peak-to-Gene Link Analysis Identifies Potential *LAMB3* Super-Enhancer and Confirms Previously Annotated *CD47* Super-Enhancer



Supplementary Figure S1. Quality metrics for scRNA-seq and scATAC-seq.

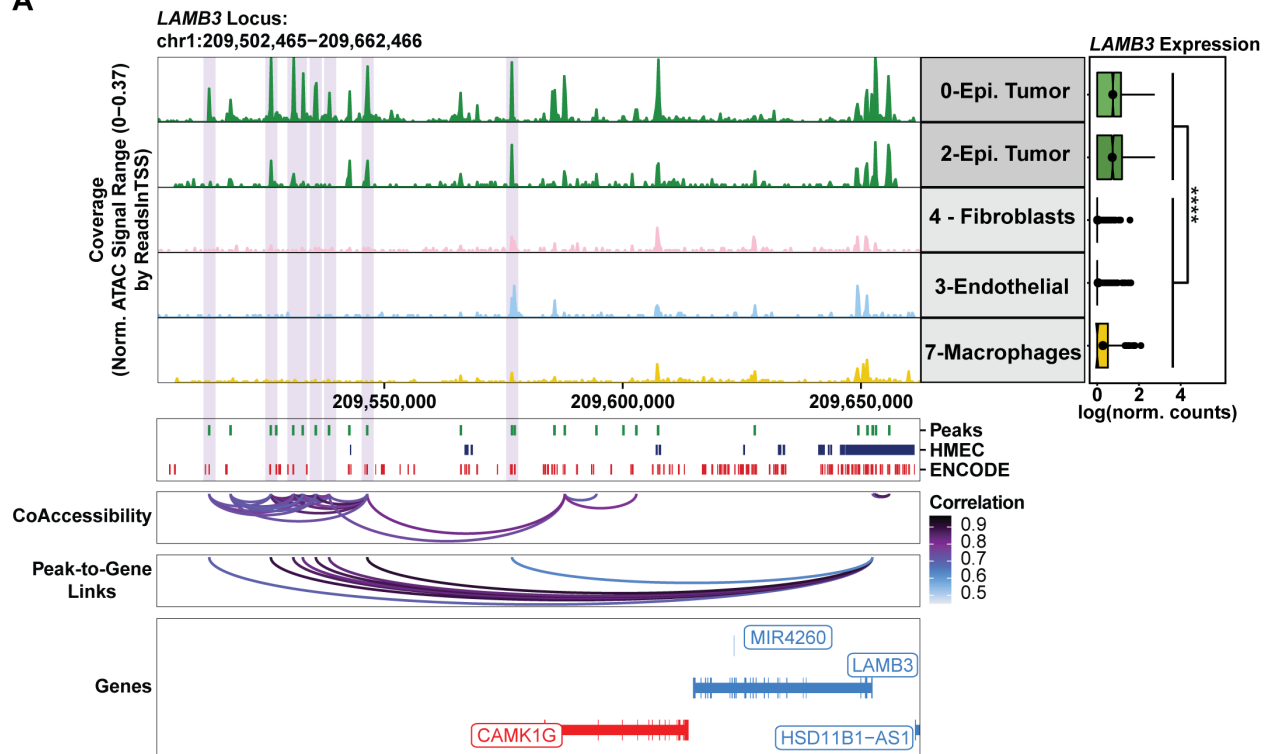
- A. scRNA-seq UMAP plot colored by cell-type, as in Fig. 1B, and numbered by cell-type cluster.
- B. Bar chart showing the number of scRNA-seq cells in each cell-type cluster.
- C. Bar chart showing the mean number of RNA counts per cell-type cluster. Red line denotes the cutoff threshold of 5,000 RNA counts.
- D. Bar chart showing the mean number of RNA features per cell-type cluster. Red line denotes cutoff threshold of 2,000 RNA features.
- E. scATAC-seq UMAP plot colored by cell type and numbered by cell-type cluster prior to filtering out low quality cells. Three cell-type clusters 1-Epi. Unassigned, 8-Epi. Unassigned, and 6-Epi. Tumor were filtered out in the UMAP in Fig. 1B.
- F. Bar chart showing the number of scATAC-seq cells in each cell-type cluster. The red line denotes the cutoff threshold of 500 ATAC cells.
- G. Bar chart showing the mean number of fragments per scATAC-seq cell-type cluster. Red line denotes the cutoff threshold of 5,000 ATAC fragments.
- H. Bar chart showing the mean TSS enrichment per scATAC-seq cell-type cluster.



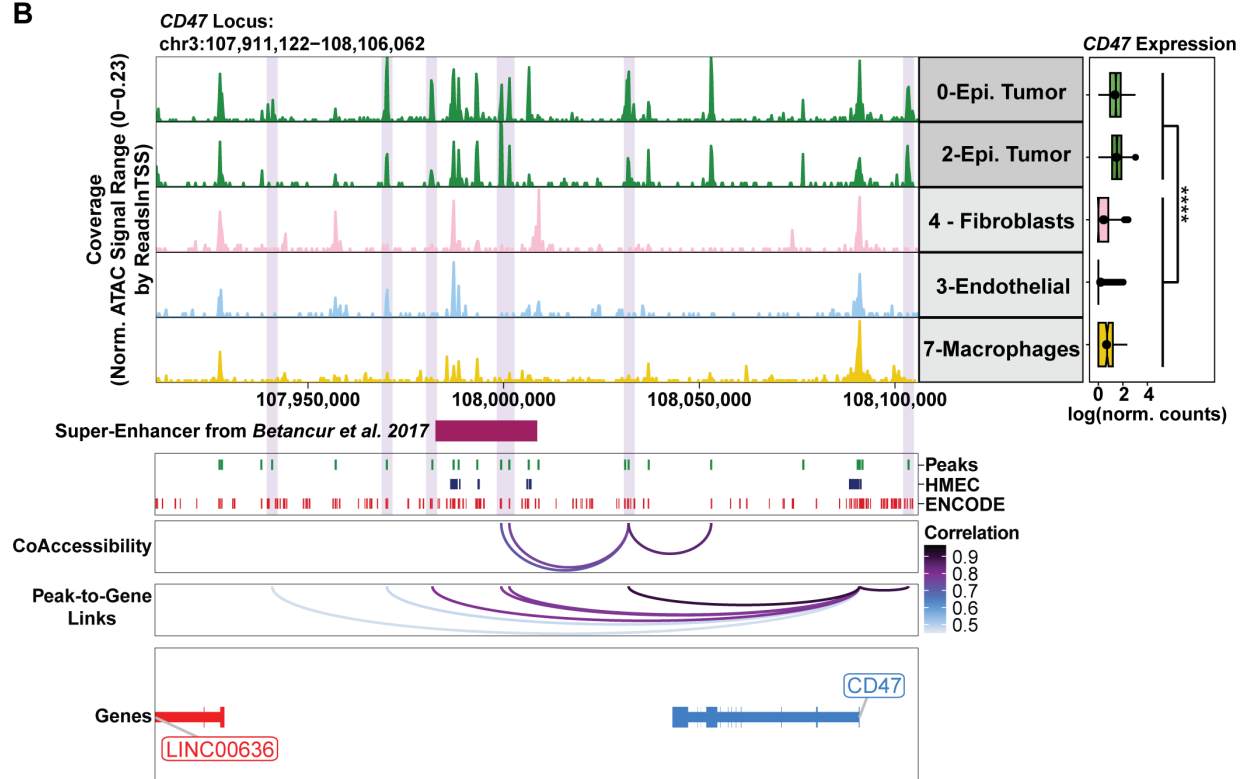
Supplementary Figure S2. CNV Thresholds for scRNA-seq Cells and QC Thresholds for scATAC-seq Cells

- scRNA-seq UMAP plot colored by CNV value.
- Scatterplot of Patient 1 inferCNV values vs. correlation. Correlation is measured as Kendall Tau with top 5% CN altered cells. Cells in red are considered cancer cells, cells in grey are considered unassigned cells, and cells in blue are considered normal cells.
- Scatterplot of Patient 2 inferCNV values vs. correlation. Correlation is measured as Kendall Tau with top 5% CN altered cells. Cells in red are considered cancer cells, cells in grey are considered unassigned cells, and cells in blue are considered normal cells.
- Scatterplot of barcode QC for Patient 1 scATAC-seq based on thresholds of TSS enrichment and log₁₀(unique fragments).
- Scatterplot of barcode QC for Patient 2 scATAC-seq based on thresholds of TSS enrichment and log₁₀(unique fragments).

A



B



Supplementary Figure S3. Peak-to-Gene Link Analysis Identifies Potential *LAMB3* Super-Enhancer and Confirms Previously Annotated *CD47* Super-Enhancer

- A. Browser track view of chromatin accessibility at the *LAMB3* locus across epithelial cancer cells (Epi. Tumor) and all other cell-type clusters. Predicted enhancers of *LAMB3* are highlighted in light purple. The peak track below the browser track denotes all scATAC-seq peaks from this study (Peaks), regulatory elements found in human mammary epithelial cells (HMEC), and all ENCODE cis-regulatory elements. The peak-to-gene correlation loops show the correlation between *LAMB3* and the peaks linked to this gene. Gene expression of *LAMB3* in matched scRNA-seq cells is depicted to the right of the browser track. **** denote p -value $< 2.22\text{e-}16$ (Wilcoxon rank-sum test).
- B. Browser track view of chromatin accessibility at the *CD47* locus across epithelial cancer cells (Epi. Tumor) and all other cell-type clusters. Predicted enhancers of *CD47* are highlighted in light purple. The *CD47* super-enhancer from Betancur *et al.*⁵⁹ is denoted in magenta. The peak track below the browser track denotes all scATAC-seq peaks from this study (Peaks), regulatory elements found in human mammary epithelial cells (HMEC), and all ENCODE cis-regulatory elements. The peak-to-gene correlation loops show the correlation between *CD47*, and the peaks linked to this gene. Gene expression of *CD47* in matched scRNA-seq cells is depicted to the right of the browser track. **** denote p -value $< 2.22\text{e-}16$ (Wilcoxon rank-sum test).