Integrated single-cell and spatial transcriptomics reveals hypoxic CAF subtype associated with lymph node metastasis in penile cancer

Feiya Yang[1,2,3,#], Haoyuan Shi[1,2,3,#], Liyuan Wu[4,#], Haoxi Wang[1,2,3,#], Yong Liu[5,#], Huaqi Yin[6], Yiwei Liu[1,2,3], Lu Yin[1,2,3], Wenkuan Wang[1,2,3], Dong Chen[1,2,3], Kehao Pan[1,2,3], Tiejun Yang[6,], Xigao Liu[7,*], Hui Xu[8,*], Mo Zhang[9,*], Nianzeng Xing[1,2,3,*]*

[1]Department of Urology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China.

[2]Beijing Key Laboratory of Urologic Cancer Cell and Gene Therapy, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China.

[3]State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China.

[4]Department of Pathology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China.

[5]Department of Urology, Weihai Municipal Hospital, Cheeloo College of Medicine, Shandong University, Weihai, 264200, China.

[6]Department of Urology, The Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital, No. 127, Dong Ming Road, Zhengzhou, 450000, China.

[7]Department of Urology, Qilu Hospital, Cheeloo College of Medicine, Shandong University, 107 West Wen Hua Road, Jinan, 250012, China.

[8]Department of Urology, Affiliated Hospital of Chengde Medical University, Chengde, 067000, China.

[9]Department of Urology, The First Hospital of China Medical University, Shenyang, China, Institute of Urology, China Medical University, Shenyang, China.

[#]These authors contributed equally to this work.

[*]Corresponding authors: xingnianzeng@126.com (N.X.); peterzhang623@gmail.com (M.Z.); chengdexuhui@163.com (H.X.); Tiejunyang@126.com (T.Y.); 201362006516@email.sdu.edu.cn (X.L.)

Abstract

Background and objective

Penile cancer (PeCa) is a rare but aggressive squamous malignancy, and lymph node metastasis is its major determinant of survival. Although cancer-associated fibroblasts (CAFs) regulate metastasis in several squamous cell carcinomas, their heterogeneity and mechanistic contribution to nodal dissemination in PeCa remain poorly defined.

Methods

We integrated single-cell RNA sequencing with Visium spatial transcriptomics to map the PeCa microenvironment and its association with lymph node metastasis. Cell-state, transcription-factor, hypoxia, spatial niche, and cell-cell communication analyses were combined with prognostic evaluation in public squamous cell carcinoma cohorts and experimental validation using multiplex immunofluorescence, co-culture, Transwell, qPCR, and western blot assays.

Key findings and limitations

Metastatic PeCa contained malignant cell states marked by enhanced stemness, hypoxia adaptation, and immune evasion, together with expansion of a hypoxia-adapted myofibroblastic CAF subset, myCAF_ACTA2. Spatial mapping localized myCAF_ACTA2 to hypoxic peritumoral niches enriched for metastasis-associated cellular communities. Cell-cell communication analysis identified POSTN-(ITGAV+ITGB5) signaling from myCAF_ACTA2 to tumor cells as a candidate pro-metastatic axis. Regulatory network analysis highlighted BHLHE41 as a key upstream transcription factor in myCAF_ACTA2, and BHLHE41 expression was positively correlated with POSTN. In vitro, BHLHE41 upregulation in CAFs increased POSTN expression and enhanced epithelial-mesenchymal transition and invasion of penile cancer cells, whereas POSTN knockdown or integrin blockade attenuated these effects. High expression of ACTA2, BHLHE41, and POSTN was associated with worse survival across squamous cell carcinoma cohorts. Limitations include the single-case spatial transcriptomic analysis, the relatively small cohort size, and the lack of in vivo validation.

Conclusions and clinical implications

These data identify a hypoxia-adapted myCAF subtype associated with lymph node metastasis in PeCa and support an in vitro mechanistic model in which the BHLHE41-POSTN axis in myCAF_ACTA2 may contribute to tumor cell epithelial-mesenchymal transition through ITGAV/ITGB5-related signaling.

Key words

Penile cancer, cancer-associated fibroblasts, single-cell RNA sequencing, spatial transcriptomics, lymph node metastasis

Main

Penile cancer (PeCa) is a rare but aggressive malignancy of the male genitourinary tract [1,2], with the vast majority of cases pathologically classified as squamous cell carcinoma (SCC) [3]. The extent of lymph node metastasis is the most important prognostic factor for survival in patients with PeCa, with pelvic or inguinal lymph node metastases being the most common sites of initial spread [4,5]. For primary tumors, radical surgery is associated with a 5-year survival rate exceeding 80% [6,7]. However, compared to patients without metastasis, patients with metastatic disease have a substantially lower 5-year survival rate of approximately 16% [8,9]. Because of psychosocial and other factors, the diagnosis of PeCa is often delayed, with up to 50% of patients experiencing a delay of over one year before receiving treatment, leading to disease progression and metastasis [10]. In China, as many as 40% of newly diagnosed PeCa cases present with lymph node metastasis [11]. In recent years, EGFR-targeted therapies have shown preliminary activity in the palliative treatment of advanced PeCa, though evidence remains limited to small case series [12-14].

In PeCa, the management of patients with lymph node metastasis has evolved to include multimodal therapy and systemic chemotherapy before surgery, with adjuvant chemotherapy potentially administered after pelvic lymph node dissection. The role of radiotherapy in neoadjuvant or adjuvant settings remains largely unclear due to the lack of high-level evidence supporting its potential benefits [15,16]. The clinical significance of molecular biomarkers and genomic profiling in PeCa requires further investigation. Although several molecular markers associated with disease progression have been identified, such as SCC antigen [17], E-cadherin [18], MMP2 and MMP9 [19], the clinical utility of these markers remains limited.

The limitations of these molecular biomarkers have prompted researchers to focus on in-depth investigations of the tumor microenvironment. Among its diverse cellular components, cancer-associated fibroblasts (CAFs) have received considerable attention because of their roles in tumor progression and metastasis. In various squamous cell carcinomas (SCCs), including head and neck, esophageal and cervical cancers, CAFs specifically modulate the differentiation state of tumor cells through the secretion of factors such as TGF-β and PDGF [20,21]. On the one hand, CAFs induce epithelial-mesenchymal transition (EMT), enabling cancer cells to acquire invasive and metastatic capabilities [22]. On the other hand, CAFs suppress normal keratinizing differentiation, promoting an undifferentiated and highly invasive tumor-cell phenotype. In addition, CAFs directly promote intratumoral lymphangiogenesis by secreting lymphangiogenic factors such as VEGF-C, VEGF-D and FGF-2, thereby facilitating lymphatic metastasis of SCC [23]. Although the mechanistic roles of CAFs have been extensively studied in other SCCs, their characteristics, functional heterogeneity and specific molecular mechanisms in regulating lymph node metastasis remain poorly understood in penile squamous cell carcinoma. This knowledge gap not only limits our understanding of the metastatic biology of PeCa but also hinders the development of microenvironment-based precision therapeutic strategies.

Single-cell sequencing and spatial transcriptomics have provided unprecedented opportunities to investigate the heterogeneity of the tumor microenvironment in depth [24]. These technologies can reveal the transcriptomic profiles, intercellular interactions and spatial distribution patterns of different cell types in tumor tissues at single-cell resolution. However, due to the rarity of PeCa, single-cell studies of its tumor microenvironment remain scarce, and the underlying mechanisms are incompletely understood.

To address these gaps, this study employed single-cell RNA sequencing and Visium spatial

transcriptomics to systematically characterize the heterogeneity of CAFs in the PeCa tumor microenvironment and elucidate their role in regulating EMT and lymph node metastasis. We identified a hypoxic myCAF subtype that was associated with lymph node metastasis in PeCa and may contribute to a pro-invasive tumor phenotype through POSTN-(ITGAV+ITGB5) intercellular signaling. This subpopulation was enriched in metastatic PeCa and localized to metastasis-associated spatial niches in the representative spatially profiled case. Our findings provide molecular insights into the metastatic biology of PeCa and a rationale for developing microenvironment-informed therapeutic strategies.

Methods

Patient sample collection

This study was reviewed and approved by the Independent Ethics Committee of the National GCP Center for Anticancer Drugs (Ethics Approval No.: NCC2024C-663). This study was conducted in accordance with the principles of the Declaration of Helsinki and ICH-GCP guidelines. All patients who provided human specimens signed a written informed consent form. All authors approved submission of the manuscript and take responsibility for the accuracy and integrity of the data.

From January 2023 to October 2024, we collected 8 specimens from patients with primary PeCa who underwent surgical resection at Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College. Additionally, one FFPE tissue sample from a treatment-naïve PeCa patient with synchronous lymph node metastasis was obtained for Visium HD spatial transcriptomic profiling.

scRNA-seq library preparation

Tissues were transported in sterile culture dishes containing 10 mL of 1 × Dulbecco's Phosphate-Buffered Saline (DPBS; Thermo Fisher, Cat. No. 14190144) on ice to remove any residual storage solution. The tissues were then minced on ice. We used a dissociation solution containing 0.25% Trypsin (Thermo Fisher, Cat. No. 25200-072) and 10 μg/mL DNase I (Sigma, Cat. No. 11284932001) dissolved in DPBS with 5% Fetal Bovine Serum (FBS; Thermo Fisher, Cat. No. SV30087.02) to digest the tissues. The tissues were dissociated at 37 °C with a shaking speed of 50 rpm for approximately 40 min. We periodically collected the dissociated cells every 20 min to enhance cell yield and viability. Cell suspensions were filtered through a 40 μm nylon cell strainer, and red blood cells were lysed using 1 × Red Blood Cell Lysis Solution (Thermo Fisher, Cat. No. 00-4333-57). The dissociated cells were then washed with 1 × DPBS containing 2% FBS. Cell viability was assessed by staining with 0.4% Trypan blue (Thermo Fisher, Cat. No. 14190144) using the Countess® II Automated Cell Counter (Thermo Fisher).

10x library construction and sequencing

Single-cell suspensions were loaded onto the 10x Genomics Chromium platform to generate Gel Beads-in-Emulsion (GEMs), with individual cells partitioned together with barcoded beads. Upon exposure to cell lysis buffer, polyadenylated RNA molecules hybridized to the beads. These beads were then collected into a single tube for reverse transcription, where each cDNA molecule was tagged at the 5′ end, corresponding to the 3′ end of a messenger RNA transcript, with a UMI and a cell barcode identifying its cell of origin. Subsequently, the beads underwent second-strand cDNA synthesis, adapter ligation, and universal amplification. Sequencing libraries were prepared from fragmented cDNA products to enrich for the 3' ends of transcripts associated with the cell barcode and UMI. All remaining procedures, including library construction, were conducted

according to the standard protocol provided by the manufacturer (CG000206 Rev D). The sequencing libraries were quantified using a High Sensitivity DNA Chip (Agilent) on a Bioanalyzer 2100 and the Qubit High Sensitivity DNA Assay (Thermo Fisher Scientific). The libraries were then sequenced on a NovaSeq 6000 (Illumina) using paired-end 150-bp sequencing.

Visium HD Spatial sequencing

Formalin-fixed, paraffin-embedded (FFPE) samples that passed RNA quality control (DV200 > 30%) were used for spatial transcriptomic analysis and sequencing. Sections measuring 5 μm in thickness were mounted onto Visium Gene Expression slides (10x Genomics), baked at 42 °C for 3 h, and dried overnight at room temperature in a desiccator. For deparaffinization, slides were incubated at 60 °C for 2 h, immersed in xylene, and rehydrated through an ethanol gradient. Hematoxylin and eosin (H&E) staining was performed using Mayer's hematoxylin (Millipore Sigma), bluing reagent (Dako, Agilent), and alcoholic eosin (Millipore Sigma). After staining, slides were scanned under a microscope. Probe hybridization, probe release, and library construction were performed using the Visium HD Spatial Gene Expression Reagent kit from 10x Genomics (PN-1000675 for Human, 6.5 mm) according to the User Guide (CG000685). Transcriptome libraries were subjected to high-throughput sequencing using PE-150 mode.

scRNA-seq data pre-processing

The newly generated scRNA-seq data from our cohort were processed using the Cell Ranger pipeline (Version 7.1.0), with gene count matrices generated using the default and recommended parameters. In addition, publicly available scRNA-seq data from 16 treatment-naïve PeCa patients reported in a previous study were downloaded and integrated with our dataset. The final integrated dataset consisted of 24 treatment-naïve PeCa patients. After quality control, batch correction and downstream analyses were performed on the integrated dataset [25]. FASTQ reads were aligned to the GRCh38 reference genome using the STAR algorithm. We used the Seurat R package (Version 5.0.1) for all subsequent analyses. To eliminate the influence of low-quality cells such as empty droplets and multiplets, cells with fewer than 200 or more than 6,000 detected genes were excluded. Cells with mitochondrial UMI fractions ≥ 25% were excluded.

We used the Seurat::NormalizeData, Seurat::FindVariableFeatures, and Seurat::ScaleData functions to normalize and scale the data. The Seurat::IntegrateLayers function was used to reduce technical variability. The top 50 principal components (PCs) were calculated using the Seurat::RunPCA function. UMAP dimensionality reduction was then performed using the top 50 PCs to obtain a two-dimensional representation.

Cell cluster and annotation

We conducted an initial clustering analysis on the integrated data using Seurat::FindNeighbors and Seurat::FindClusters (resolution = 0.8). The resulting cell clusters were manually annotated based on well-known marker genes. Marker genes included T cells (CD3D, CD3E, CD3G), B cells & plasma cells (MS4A1, CD19, CD79A, CD79B, IGHM), myeloid cells (CD14, CD68, CD163, LYZ, ITGAX), endothelial cells (PECAM1, VWF, ENG, CDH5, MCAM), stromal cells (COL1A1, COL3A1, FN1, DCN, LUM), and epithelial cells (EPCAM, KRT14, KRT17, KRT19, CDH1, MUC1). Visualization was performed using the Seurat::DotPlot function, and cell identities were annotated based on the expression levels and proportions of marker genes within the clusters.

Next, we performed a second round of clustering to further characterize the subtypes of major cell types within the tumor microenvironment (TME). To mitigate potential data heterogeneity caused by cell types, cell numbers, or inter-sample variability, we first filtered out features with an

expression proportion below 50% within the cell type in the second round of clustering and used the Harmony method for batch correction. In the second clustering, we used 50 principal components (PCs) for dimensionality reduction and unsupervised clustering (resolution = 1).

To support cell-subtype annotation, we used multiple large language models (LLMs) as an auxiliary annotation strategy. Cluster-specific marker genes were identified using Seurat::FindAllMarkers with only.pos = TRUE, min.pct = 0.25, and logfc.threshold = 0.5. The marker-gene lists are provided in Supplementary Table S2. The top marker genes from each cluster were supplied to Claude-3.5-Sonnet, Gemini-2.0-Flash, Qwen-Max, and GPT-4o to obtain preliminary annotation suggestions. The complete annotation prompt is provided in the Supplementary Methods.

The prompt instructed the models to assign cell identities solely from the supplied marker genes, using standard nomenclature from the Cell Ontology and established single-cell literature. Unsupported tissue assumptions and overly granular subtype labels were discouraged. For each cluster, the requested outputs included a primary predicted cell type, an integer confidence score from 0 to 100, a brief justification identifying supporting and contradictory markers, one or two alternative candidates when the profile was ambiguous, and caveats concerning missing canonical or conflicting markers. The specified confidence categories were high confidence (90-100), moderate confidence (70-89), and low confidence (< 70). The models were instructed to return "unassigned" when the marker combination did not support a credible cell identity. These confidence scores represent model-generated assessments and should not be interpreted as independently validated probabilities of annotation accuracy.

Final cell subtype identities were determined through expert manual curation integrating the LLM-generated suggestions, canonical marker expression, and previously published single-cell studies. Manual review assessed whether the proposed identity was consistent with the major lineage markers, whether subtype-specific markers supported the proposed level of annotation, and whether the expression profile was compatible with the reported biological characteristics of that population. Marker expression levels and the proportions of cells expressing the markers were examined using Seurat::DotPlot.

Prompt used for LLM-assisted cell subtype annotation

You are an expert computational biologist specializing in single-cell RNA-seq, cell ontology, and cell type annotation. Your task is to annotate a cell cluster purely based on its top marker genes.

Rules you must strictly follow:

1. Use standard, widely adopted cell type nomenclature from the Cell Ontology and major single-cell publications in the relevant tissue. Do NOT create novel cell names or overly granular subpopulation labels unless strongly supported by the marker genes.
2. For the given marker gene set, output:
   (a) Primary predicted cell type
   (b) Confidence score (integer 0–100): 90–100 = high confidence; 70–89 = moderate confidence; < 70 = low confidence
   (c) Brief justification: explain which markers support this cell type and whether any markers contradict this assignment
   (d) 1–2 alternative candidate cell types if the marker profile is ambiguous
   (e) Caveats: note missing canonical markers or conflicting markers that limit annotation confidence

3. Do NOT make assumptions beyond the provided marker genes. Do NOT infer tissue origin unless explicitly stated. Do NOT rely on prior knowledge of cluster size, UMAP position, or other external information. Your annotation must be driven only by the marker gene list provided.
4. If no credible cell type can be assigned from the markers, assign the label "unassigned" and clearly explain why the marker combination is insufficient for reliable annotation.
5. Keep your answer concise and structured. Avoid long paragraphs.

RCTD and BANKSY analyses

Using the scRNA-seq dataset as the reference, spacexr (version 2.2.0) was applied to infer the cell-type composition of each bin. Specifically, the default parameters were used in the create.RCTD function, requiring > 1 cell per cell type and > 1 unique molecular identifier (UMI) count per pixel; the mode parameter was set to "doublet" in the run.RCTD function. BANKSY incorporated information from neighboring spots to improve spatial clustering (lambda parameter set to 0.2). BANKSY was run using the RunBanksy function from the SeuratWrappers library. For niche composition analysis, enrichment of each cell type within individual niches was assessed using a one-sided Wilcoxon rank-sum test, with adjusted $p < 0.05$ considered statistically significant.

Cell-cell communication analysis

Cell-cell communication analysis was performed on the scRNA-seq data using CellChat (Version 2.1.2). Specifically, only receptors and ligands expressed in > 1% of cells were further evaluated, and no interaction was inferred when either the ligand or receptor was undetectable.

Pathway enrichment analyses

Gene Ontology (GO) enrichment analysis. To identify enriched Gene Ontology (GO) biological processes based on differentially expressed genes (DEGs), DEGs from each cell type were analyzed using the clusterProfiler::gseGO function in the clusterProfiler R package (Version 4.8.1), with the org.Hs.eg.db annotation package (Version 3.17.0). Gene set enrichment analysis (GSEA). GSEA was performed using the GseaVis::gseaNb function in the GseaVis R package (Version 0.0.8) (https://github.com/junjunlab/). Gene lists preranked by avg_log2FC values calculated using Seurat::FindAllMarkers were used as input. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed using the ClusterGVis::enrichCluster function in the ClusterGVis R package (Version 0.1.0), with the org.Hs.eg.db annotation package (Version 3.17.0). Single-sample GSEA (ssGSEA) was performed using the singleseqgset R package (Version 0.1.2.9000). We used the human hallmark gene sets from the msigdbr package (Version 7.5.1) as the gene sets of interest for enrichment analysis. Results with $p < 0.05$ were considered statistically significant.

CytoTRACE analysis

We inferred the developmental trajectory of human tumor cells using CytoTRACE (v.0.3.3). CytoTRACE is based on the concept of transcriptional diversity (i.e., the number of genes expressed in cells decreases during differentiation). A $\log_2$-normalized expression matrix was used as input.

Copy-number variation inference

To further evaluate malignant epithelial cells, we inferred copy-number variation (CNV) profiles using the R package inferCNV, with immune, stromal, and endothelial cells as reference populations and epithelial cells as the test population.

TCGA data and Survival analysis

We used the R package easyTCGA (Version 0.0.4.200.0) to download the TCGA pan-cancer gene

expression matrix and clinical information. Kaplan-Meier survival curves were plotted using ggsurvplot (Version 0.4.9) function in the R package Survminer (Version 0.4.9). Cox proportional hazards models were used to estimate hazard ratios and 95% confidence intervals.

Functional module scoring

Module scores in the scRNA-seq and Visium HD datasets were calculated using the AddModuleScore function in Seurat, with genes from the hypoxia gene set (HARRIS_HYPOXIA from MSigDB).

Regulon network

Regulatory-network analysis was performed using the R package SCENIC (v.1.1.3) to assess coexpression between transcription factors and their putative target genes. We built and scored a gene regulatory network using the default parameters. The raw count matrix was used to build a coexpression network using the runCorrelation and runGENIE3 functions, and regulon activity was evaluated with AUCell. Regulon activity for each cell was calculated as the average normalized expression of putative target genes. To identify differentially active transcription factors between cells of interest and other cell types, we extracted regulonAUC data from SCENIC results and calculated mean activity scores for each group. Transcription factors with higher activity in cells of interest were identified using fold change and difference thresholds (Fold_Change > 75th percentile, Difference > 75th percentile, Mean activity > 0.005).

Primary PeCa cell and CAF isolation and culture

Primary PeCa cell and cancer-associated fibroblast (CAF) isolation was attempted using all available freshly resected penile cancer specimens. Paired primary PeCa cells and CAFs could be successfully established and expanded from three independent patient specimens and were used for the functional experiments. Because this study focused on lymph node metastasis, most available specimens were obtained from patients with T3N1M0 or T3N2M0 disease and were HPV-negative according to the clinical pathology reports. The three independent patient-derived PeCa cell and CAF cultures were used as biological replicates for the co-culture, Transwell, qPCR, and Western blot experiments.

Fresh tumor tissue was transferred immediately into precooled sterile PBS containing penicillin-streptomycin and washed three times. Visible adipose tissue, blood vessels, necrotic tissue, and excess connective tissue were removed, and the remaining viable tumor tissue was washed twice with PBS. The tissue was then finely minced with sterile ophthalmic scissors and transferred to a 15 mL centrifuge tube for enzymatic digestion with type II collagenase and 0.25% trypsin. Digestion was performed at 37 °C for 20-30 min with gentle mixing every 5 min. The digestion was terminated by adding an equal volume of complete medium containing 10% fetal bovine serum.

The resulting cell suspension was gently dispersed and filtered through a 70 μm sterile cell strainer to remove undigested tissue fragments. Cells were collected by centrifugation at 1000 r/min for 5 min, washed once with prewarmed PBS, and resuspended in DMEM/F12 complete medium. Differential adhesion was then used to enrich the epithelial tumor-cell and fibroblast populations. The mixed cell suspension was plated and incubated at 37 °C with 5% CO_2 for 2 h. Fibroblast-like stromal cells preferentially adhered during this period, whereas the nonadherent cell fraction was transferred to a new culture dish for further culture of PeCa cells. The adherent fibroblast-enriched population was maintained separately as the CAF culture.

Cell morphology, adherence, and growth status were dynamically monitored during primary culture. The culture medium was first replaced after 24 h to remove dead and nonadherent cells and was

subsequently changed every 48 h. Cultures showing stable adherent growth, relatively uniform morphology, and no evidence of microbial contamination were expanded for subsequent experiments. CAF identity was further supported by fibroblast morphology and immunofluorescence staining for ACTA2 and COL1A1.

Multiplex Immunofluorescence Staining

Immunofluorescence staining of cultured PeCa-associated fibroblasts (CAFs) was performed on coverslips. Multiplex immunofluorescence was performed on PeCa tissue sections obtained at our center. When the cells reached the exponential growth phase, they were digested with trypsin, counted, and seeded at a density of 1×10^5 cells per well into 12-well plates pre-placed with cell climbing slices, followed by incubation at 37 °C in a 5% CO_2 incubator for 24 h. After incubation, the culture medium was discarded, and the cells were washed three times with pre-warmed 1 × PBS. Subsequently, 4% paraformaldehyde was added for fixation at room temperature for 20 min. After fixation, the cells were washed again three times with 1 × PBS. Permeabilization was performed using 0.2% Triton X-100 at room temperature for 2-5 min, followed by three washes with 1 × PBS (10 min each). Blocking was carried out with 5% BSA at room temperature for 30 min. Primary antibodies diluted in 1% BSA were then added, and the samples were incubated overnight at 4 °C in a humidified chamber. The following day, after washing three times with 1 × PBS, fluorescent secondary antibodies diluted in 1% BSA were applied and incubated at room temperature for 30 min in the dark. The samples were washed three more times with 1 × PBS and finally mounted using an anti-fade mounting medium containing DAPI. Cells were observed and imaged using an inverted fluorescence microscope. Cell culture dishes (Cat. No. 430167), 12-well culture plates (Costar, Cat. No. 3512), and 12-well cell climbing slices (WHB, Cat. No. WHB-12-CS) were used. Fluorescence images were acquired using an Olympus CKX53 inverted fluorescence microscope equipped with a DP70 camera.

Quantitative real-time polymerase chain reaction

Following co-culture, total RNA was extracted from the designated cell population using TRIzol. RNA concentration and purity were assessed with a nucleic acid analyzer using the following quality criteria: an A260/A280 ratio between 1.8 and 2.0, an A260/A230 ratio ≥ 2.0, and a concentration ≥ 50 ng/μL, before downstream analysis. Following the manufacturer's instructions, total RNA was reverse transcribed into cDNA using a reverse transcription kit. The reverse-transcription reaction mixture contained the RNA template, reverse transcriptase, primers, dNTPs, and the remaining kit components in a total volume of 20 μL. The thermal cycling protocol was set as follows: 25 °C for 5 min, 42 °C for 30–60 min, and 85 °C for 5 min. Upon completion, the obtained cDNA was diluted with sterile, nuclease-free water, aliquoted, and stored at -20 °C for later use.

Subsequently, quantitative real-time PCR (qPCR) was performed using the SYBR Green method. The qPCR reaction mixture consisted of 10 μL of qPCR Master Mix, 1.0 μL each of the forward and reverse primers, 3 μL of cDNA template, and sufficient sterile, nuclease-free water to reach the final reaction volume. qPCR was performed using the following cycling conditions: initial denaturation at 95 °C for 30 s; followed by 40 cycles of amplification (denaturation at 95 °C for 5 s, annealing/extension at 60 °C for 30 s). Melting-curve analysis was performed after the amplification cycles to verify the specificity of the amplification products.

Relative target-gene expression was calculated using the $2^{\wedge(-\Delta\Delta CT)}$ method. Ct values were normalized to β-actin to obtain ΔCt values, and ΔΔCt values were calculated relative to the control group to determine fold changes in gene expression.

Western blot

Cultured cells were washed with precooled PBS and lysed on ice for 20-30 min in RIPA buffer containing PMSF. Protein concentrations were then determined using the BCA assay: a standard curve was prepared, and the samples were incubated with the BCA working solution at 37 °C for 20-30 min. Absorbance at 562 nm was measured to calculate the protein concentration of each sample. Protein samples were mixed with 5 × loading buffer and heat-denatured. An SDS-PAGE gel of appropriate concentration was prepared according to the molecular weight of the target protein, and 30-60 μg of protein was loaded per lane. Electrophoretic separation was performed using the separating gel under a constant voltage of 120 V. After electrophoresis, proteins were transferred to a PVDF membrane activated with methanol using the wet transfer method. Proteins were transferred at 300 mA for 30–60 min. The membrane was then rinsed with TBST and blocked with 5% skim milk at room temperature for 1 h. After blocking, the membrane was incubated with specific primary antibodies at room temperature for 2 h, followed by three washes with TBST. It was then incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies at room temperature for 2 h and thoroughly washed. Finally, the membrane was incubated with ECL chemiluminescent substrate, and chemiluminescent signals were captured using a Tanon-4600 imaging system. Band intensities for the target proteins and the loading control (β-actin) were quantified using image-analysis software. Relative protein expression was calculated by normalizing target-band intensity to β-actin for between-group comparisons. The BCA protein assay kit used for protein quantification was Cat. No. DXWB010. Primary antibodies included POSTN (Proteintech, Cat. No. 66491-1-Ig, 1:5000), N-cadherin (Affinity, Cat. No. AF4039, 1:1000), E-cadherin (Affinity, Cat. No. AF0131, 1:1000), Vimentin (Affinity, Cat. No. AF7013, 1:1000), and β-actin (Proteintech, Cat. No. 66009-1-Ig, 1:2000). HRP-conjugated secondary antibodies were used at a dilution of 1:5000.

Western blot experiments were performed in three independent batches. To evaluate the relationship between BHLHE41 and POSTN, the following experimental groups were established: a PeCa and CAF co-culture group, a PeCa and BHLHE41-overexpressing CAF co-culture negative control group, a PeCa and BHLHE41-overexpressing CAF co-culture group, a PeCa and si-BHLHE41 CAF co-culture negative control group, and a PeCa and si-BHLHE41 CAF co-culture group. After co-culture, CAFs were separately lysed and used for Western blot analysis of POSTN expression. In parallel, PeCa cells from the corresponding co-culture conditions were separately harvested and used for Western blot analysis of the EMT markers E-cadherin, N-cadherin, and Vimentin. Thus, POSTN expression was assessed in CAFs, whereas EMT-marker expression was assessed specifically in PeCa cells.

RGD peptide blockade

To evaluate the functional involvement of POSTN-related integrin signaling in the co-culture system, an integrin-blocking Arg-Gly-Asp (RGD) peptide was used. Briefly, PeCa cells were co-cultured with CAFs in the presence of RGD peptide (MCE, HY-P1740) at a final concentration of 10 μM for 12 h. PeCa cells co-cultured with untreated CAFs under the same conditions were used as the control group. After treatment, PeCa cells were separately harvested for total RNA extraction and subsequent qPCR analysis, while parallel PeCa-cell samples were lysed for Western blot analysis of EMT markers. This experiment was performed to assess whether blockade of integrin-related signaling could attenuate POSTN-associated EMT marker changes in PeCa cells.

Transwell assay

Transwell assays were performed using PeCa cells. Corning Transwell inserts (product number 3422,

8 μm pore size) were prewetted in 1 × PBS for 5 min in 24-well plates. Cells were digested with trypsin, resuspended in complete medium, and adjusted to a density of 2×10^4 cells/mL. A 0.1 mL aliquot of the cell suspension from each group was seeded into the upper chamber, while 0.6 mL of complete medium was added to the lower chamber. The culture plate was then incubated in a 37 °C, 5% CO_2 incubator for 16 h. After incubation, 1 mL of 4% paraformaldehyde solution was added to each well to fix the cells at room temperature for 15 min. The fixative was removed, and the cells were washed twice with 1 × PBS. Subsequently, 1 mL of 0.5% crystal violet solution (Sinopharm, product number 71012314) was added to each well and incubated for 20 min for staining. After staining, the cells were washed three times with 1 × PBS and air-dried. Finally, non-migrated cells on the upper side of the insert were carefully removed using a cotton swab. The inserts were imaged under an optical microscope (OLYMPUS, CX31), and cells that had migrated through the membrane were counted.

To evaluate PeCa-cell invasion, five co-culture systems were established: a PeCa cell and CAF co-culture group, a PeCa cell and BHLHE41-overexpressing CAF co-culture negative control group, a PeCa cell and BHLHE41-overexpressing CAF co-culture group, a PeCa cell and si-BHLHE41 CAF co-culture negative control group, and a PeCa cell and si-BHLHE41 CAF co-culture group.

Statistical analysis

For comparisons of cell-type and cell-subtype abundance between the LN+ and LN- groups, proportions were first calculated for each patient, and each patient was treated as the independent unit of analysis. Subtype enrichment was additionally summarized using the ratio of observed to expected cell numbers (Ro/e) and odds ratios (ORs) based on patient-level cell composition. For niche composition analysis, enrichment of each cell type within individual niches was assessed using a one-sided Wilcoxon rank-sum test, with adjusted $p < 0.05$ considered statistically significant. Kaplan–Meier curves were used for exploratory survival analyses, and Cox proportional hazards models were used to estimate HRs and 95% CIs. For qPCR, three technical replicate wells were averaged before ΔCt/ΔΔCt calculations and were not treated as independent biological replicates. The functional experiments were independently performed using primary PeCa cells and CAFs established from three different patient donors. Accordingly, the three patient-derived cultures were treated as independent biological replicates, while technical replicate wells or repeated measurements within the same donor were not considered independent biological replicates.

Results

Multi-omics analysis of penile cancer tissues with and without lymph node metastasis at initial presentation.

To investigate the microenvironmental remodeling features and spatial organization characteristics associated with lymph node metastasis in PeCa, we collected 8 newly generated scRNA-seq samples and integrated them with 16 publicly available PeCa scRNA-seq samples [25], resulting in a combined cohort of 24 treatment-naïve PeCa patients, including 9 patients with synchronous inguinal lymph node metastasis (LN+) and 15 patients without detectable lymph node metastasis (LN-) at initial presentation (Table S1). After quality filtering, we retained 147,660 high-quality single-cell transcriptomes (Figure 1A). Initial data processing involved batch effect correction and integration using the Harmony algorithm, followed by unsupervised clustering (Figure S1A-B). Marker-gene analysis identified eight major cell types across all patients, including B & Plasma cells (N = 7,446), CD4 T cells (N = 25,760), CD8 T cells (N = 16,235), Innate Lymphoid cells (ILCs,

N = 5,003), Myeloid cells (N = 29,705), Stromal cells (N = 34,726), Endothelial cells (N = 6,292), and Epithelial cells (N = 22,493) (Figure 1B, S1M-Q). Subsequently, we further refined the clustering within each major cell type and identified 48 unique cell subtypes. Copy number variation (CNV) analysis performed on epithelial subclusters using inferCNV revealed widespread genomic alterations among epithelial populations, supporting their malignant characteristics (Figure S1C). Based on the combined evidence of CNV profiles and epithelial marker expression, these epithelial clusters were considered tumor-enriched epithelial populations for subsequent analyses. We also identified a minor neuronal population (S09_Neuron_CDH19, N = 233), likely reflecting the low capture efficiency of neurons in scRNA-seq. To obtain an exploratory view of the spatial organization and cellular community composition associated with PeCa lymph node metastasis, we performed 10x Visium HD spatial transcriptomics on matched primary and metastatic Formalin Fixed Paraffin Embedded (FFPE) specimens from one case. In total, we obtained transcriptome data from 263,010 spots, with a median depth of 733.9 unique molecular identifiers (UMIs) per spot and 538.0 genes per spot. After quality control using Building Aggregates with a Neighborhood Kernel and Spatial Yardstick (BANKSY) for cross-sample normalization, we performed final clustering in the integrated space using Harmony (Figure S1D) [26]. To reveal cellular composition within each spatial transcriptomic spot, we employed Robust Cell Type Decomposition (RCTD) to map the scRNA-seq profiles to the Visium HD spatial transcriptome data (Figure 1C, S1E). Patient-level analysis of the scRNA-seq data showed that LN+ patients exhibited significantly reduced proportions of B and plasma cells ($p = 0.028$) and CD4 T cells ($p = 0.024$), and increased stromal-cell abundance ($p = 0.035$) in the tumor microenvironment compared to LN- patients (Figure 1D). In the LN+ group, we observed a reduction in the CD4T01_Tn_NR4A1 and CD4T05_Tem_GZMK subsets, concomitant with an expansion of CD4T06_Treg_FOXP3 and CD4T04_Tm_ANXA1 populations within the CD4 T cells (Figure 1E). The B & Plasma cells showed modest reductions (< 0.1) in B01_Bn_NR4A2 and B06_Plasma_IGHG2 subsets (Figure 1F). Among stromal cells, the proportion of S06_iCAF_IL6 was decreased, while the proportions of S05_SMC_TNNT2 and S07_myCAF_ACTA2 were increased (Figure 1G). These findings suggest potential involvement of immunosuppressive CD4 T cell subsets and S07_myCAF_ACTA2 in early lymphatic dissemination of PeCa.

Malignant subtypes show distinct transcriptomic features associated with lymph node metastasis.

To characterize transcriptomic features associated with lymph node metastasis at initial presentation in patients with PeCa, we corrected interpatient batch effects among 22,493 tumor epithelial cells. We conducted unsupervised clustering analysis and identified ten distinct tumor epithelial subtypes based on the harmonized dataset (Figure 2A). These nine malignant epithelial cell subtypes were distributed across all patients without evidence of patient-specific subtypes (Figure S2A). We analyzed differences in the distribution of tumor epithelial subtypes between the LN+ and LN- groups by calculating the ratio of observed to expected cell numbers (Ro/e) and odds ratio (OR) values. C6 and C7 malignant subtypes were more abundant in the LN+ group, with OR values of 16.82 and 6.91, respectively ($p < 0.05$, Figure 2B, Figure S2B).

We employed CytoTRACE to assess stemness scores across tumor cell subtypes, revealing that C6 (highest stemness) and C7 (fourth-highest stemness) malignant subtypes showed higher stemness scores. The C5 subtype, which was enriched in the LN- group, exhibited the highest degree of differentiation, while the C8 subtype had a higher stemness score (Figure 2C). Next, differential

gene expression analysis between the LN+ and LN- groups identified 3,064 upregulated and 724 downregulated genes in the LN+ group (Table S2). Pathway enrichment analysis of upregulated genes highlighted biological processes related to nervous system development including axonogenesis, regulation of neuron projection development, and positive regulation of neuron projection development (Figure 2D-E). Gene Set Enrichment Analysis (GSEA) further confirmed significant enrichment in GO_SYNAPSE and GO_NEUROGENESIS terms (Figure 2E). Notably, downregulated genes in the LN+ group were enriched in epidermis-development and MHC class II-related pathways (Figure 2D). The C6 and C7 subtypes, which were enriched in the LN+ group, exhibited reduced expression of both MHC class I and II genes (Figure 2F, Figure S2C), coupled with elevated immune checkpoint molecules (CD274, CD276, PDCD1LG2) (Figure 2G), suggesting their potential for immune evasion and contribution to immunosuppressive tumor microenvironment (TME) formation.

To investigate potential characteristics of early lymph node metastasis in PeCa, we performed transcription factor (TF) analysis by using the single-cell regulatory network inference and clustering (SCENIC). Our analysis identified preferential activation of TFs including SOX2, ERG, HIF1A, FOS, FOXQ1, and TCF4 in malignant subtypes associated with early lymph node metastasis (Figure 2H, Table S3), accompanied by elevated gene expression levels (Figure 2I, S2D-I). The co-activation of these TFs suggested a potential link between hypoxia adaptation and early lymph node metastasis in PeCa. Therefore, we evaluated the hypoxia signaling scores of malignant epithelial cells in patients from the LN+ and LN- groups (Figure 2J) as well as in metastatic and non-metastatic lesions (Figure 2K) in both the scRNA-seq and spatial datasets. Across both datasets, a stronger hypoxia-response signature in malignant epithelial cells was associated with early lymph node metastasis in PeCa. Differential gene expression analysis identified 15 significantly upregulated genes in early metastatic cases (Figure 2L), which were functionally enriched in extracellular matrix remodeling, immune regulation, and metabolic reprogramming pathways.

Cell-cell communication in the microenvironment of early lymph node metastasis in PeCa

After characterizing the differences between the LN+ and LN- groups, we further investigated their cell-cell communication networks within the TME to identify potential subtypes associated with early lymph node metastasis of PeCa. Analysis of interaction strength revealed significant reorganization of cell-cell communication networks in metastatic tissues, with the LN+ group exhibiting more extensive intercellular signaling patterns than the LN- group (Figure 3A-B, Figure S3A). Notably, we observed enhanced communication strength between stromal cells and various immune populations (including CD4 T, CD8 T, and myeloid cells), alongside substantial alterations in epithelial cell interaction patterns (Figure 3C). These changes suggest that PeCa cells may alter communication with microenvironmental components to evade immune surveillance, while specific stromal subtypes potentially facilitate metastatic progression through enhanced cellular crosstalk.

Differential ligand-receptor analysis identified changes in intercellular signaling associated with metastatic progression. CD4 T cells showed distinct changes in signaling in metastatic microenvironments, including downregulation of MHC class II and CD99 signaling pathways (Figure 3D left, Figure S3B). Similarly, CD8 T cells showed changes in signaling, with a trend toward reduced MHC class I-related molecule expression in the metastatic group (Figure 3D middle, Figure S3C). In epithelial cells, increased Notch, FN1, and SPP1 signaling suggested activation of EMT-related programs (Figure 3D right, Figure S3D).

Within the T cell lineage, several T cell subsets showed differential distribution between the LN+

and LN- groups. Among these populations, CD8T05_Tc_ISG15 exhibited a relatively high observed-to-expected enrichment ratio, while CD4T08_Th_ISG15 also showed increased representation in metastatic samples (Figure 3E) [27,28]. However, patient-level analysis did not demonstrate significant associations between these interferon-stimulated T cell subsets and LN status, suggesting that their enrichment may reflect inter-patient heterogeneity rather than a dominant driver of lymph node metastasis. Notably, among the CAF subtypes, S07_myCAF_ACTA2 showed marked enrichment during metastasis (OR = 6.27). Recent pan-cancer spatial multi-omics analyses have shown that distinct CAF states, including myofibroblastic populations, are associated with specific cellular neighborhoods and patterns of tumor-stromal and stromal-immune interactions [22]. S07_myCAF_ACTA2 cells showed features associated with extracellular-matrix remodeling, vascular regulation, and paracrine support of tumor-cell invasion, suggesting a potential role in stromal changes associated with early lymph node metastasis in PeCa (Figure 3E-F) [29].

Immunosuppressive molecule profiling further delineated the immunomodulatory roles of distinct stromal subtypes (Figure 3G). Differential expression patterns of key immune checkpoints and coinhibitory molecules were observed across cell types, with CD276 (B7-H3) showing the highest expression in S07_myCAF_ACTA2, suggesting an immunosuppressive function for this subtype in PeCa. The heterogeneous expression patterns of BTNL3, SLAMF7, and PDCD1LG2 (PD-L2) further indicated subtype-specific immune regulatory states that may collectively influence the balance between immune activation and immunosuppression within metastatic microenvironments.

Cross-tissue cellular modules and spatial patterns in penile carcinoma

Given the critical role of cellular niche composition in cancer progression and metastasis [30], we systematically investigated the co-occurrence patterns of cellular subtypes across distinct tissue compartments. By integrating specific marker expression of cell types with spatial proximity data, we applied an unsupervised clustering algorithm to identify 10 microenvironmental niches with distinct cellular compositions (Figure 4A). Heatmap analysis showed that each niche had a distinct pattern of subtype enrichment, consistent with functional specialization. Uniform Manifold Approximation and Projection (UMAP) visualization showed distinct distributions of tumor-microenvironment and malignant-cell niches in the integrated space (Figure 4B). We observed pronounced transcriptomic segregation between TME niches and cancer cell niches, suggesting spatial compartmentalization and functional specialization within the tumor tissue architectures.

Within this single matched case, spatial transcriptomics showed that several malignant niches were represented in both the primary and metastatic lesions (Figure 4C) and revealed distinct niche composition patterns between the primary tumor and lymph node metastatic regions within this case (Figure 4D). Lymph node tissues showed specific enrichment of niche_2 and niche_4, whereas niche_5-8 and niche_10 were predominantly localized in primary tumors, consistent with site-specific microenvironmental differences. Crucially, we found that niche_1, niche_3, and niche_9 showed a high degree of co-occurrence in both primary lesions and lymph node metastatic lesions, suggesting that these niches may represent candidate microenvironmental modules associated with lymph node metastasis in PeCa.

Analysis of cellular composition revealed that niche_1 was predominantly enriched for angiogenesis-related subtypes, including S01_Pericyte_CD36, which mediated vascular stability [31]; S02_SMC_RERGL, which maintained vascular architecture [32]; and E02_Vein_ACKR1, which facilitated tumor intravasation [33], collectively providing structural support for entry into

the circulation. Niche_2 was primarily associated with immune-cell composition and regulation and may modulate metastatic competence through altered immune responses. Niche_9 contained cell states associated with stromal remodeling and immunosuppression, including S07_myCAF_ACTA2 with strong contractile capacity, pro-tumoral M11_Mph_ISG15 that exhibited M2-like polarization characteristics and S08_apCAF_CD74 involved in the regulation of antigen presentation. Mechanistically, S07_myCAF_ACTA2 showed transcriptional features consistent with ACTA2-mediated contractility, extracellular-matrix remodeling, and growth-factor signaling that could support tumor-cell invasion and EMT. Similar pro-metastatic properties of myofibroblastic CAFs have been experimentally demonstrated in pancreatic ductal adenocarcinoma, where EGFR-activated myCAFs promoted tumor invasion and metastasis [34]. These observations raise the possibility of cooperative interactions with M11_Mph_ISG15 in PeCa.

Hypoxic myCAF subpopulations are associated with tumor progression and early metastasis in penile carcinoma.

To investigate the potential dynamic relationship among CAF subpopulations, we performed Monocle3-based pseudotime trajectory analysis using patient-aligned CAF populations. The trajectory analysis revealed a continuous CAF state manifold connecting resident fibroblast-like cells, iCAF, and myCAF_ACTA2 populations (Figure S3E-J). Along pseudotime, resident fibroblast-associated markers gradually decreased, whereas myofibroblastic and extracellular matrix remodeling programs, including ACTA2, TAGLN, POSTN, and MMP-related signatures, showed progressively increased activity. These findings suggest that S07_myCAF_ACTA2 represents an activated CAF state emerging during fibroblast remodeling rather than an entirely independent CAF lineage (Figure S3E-J). Given the prominent enrichment of S07_myCAF_ACTA2 in metastasis-associated niche_9, we further characterized the molecular features and functional programs of this CAF subset. Transcriptomic profiling revealed that S07_myCAF_ACTA2 exhibited elevated expression of functionally relevant genes, including ECM remodeling components (LUM, DCN, PDGFRA), vascular regulators (THBS2, NRP2), signaling mediators (FOXO1, INHBA, GJA1, ADAM12, NOX4, ISLR) and immune checkpoint molecules (CD276, BTN3A1, BTN3A2) (Figure 5A). This diverse expression profile was consistent with a role for this subpopulation in tumor-microenvironment remodeling.

Enrichment analysis further characterized the functional features of S07_myCAF_ACTA2, including collagen catabolism, vascular development, immune tolerance induction, and negative regulation of immune responses (Figure 5B). GSEA revealed significant enrichment of vascular development (NES = 2.501144) and tolerance induction (NES = 4.063773) gene sets, demonstrating this subtype's dual mechanistic role in coordinately establishing a pro-metastatic niche through synergistic promotion of angiogenesis and immune tolerance.

Spatial analysis in the representative Visium HD case suggested distinct distribution patterns of CAF subtypes. At the tissue level, candidate metastasis-associated niches (niche_1, niche_3 and niche_9) showed characteristic spatial patterns within this case (Figure 5C). Notably, S07_myCAF_ACTA2 showed predominant localization to peritumoral regions in the analyzed spatial sample, while S08_apCAF_CD74 showed a more distal distribution relative to tumor regions (Figures 5D-E), suggesting functional compartmentalization among CAF subtypes. Hypoxia scoring within the analyzed spatial transcriptomic dataset indicated that niche_9 exhibited relatively higher hypoxia scores, and S07_myCAF_ACTA2 similarly showed a strong hypoxia-associated signature (Figure 5F). Together with the scRNA-seq and functional analyses, these exploratory

spatial observations suggest that S07_myCAF_ACTA2 may be associated with a hypoxic, pro-invasive microenvironment. The angiogenic regulators highly expressed by this subtype may become activated under hypoxic conditions, thereby promoting aberrant angiogenesis and creating routes that facilitate tumor dissemination, ultimately contributing to microenvironmental remodeling during early lymph node metastasis in PeCa.

To explore the potential prognostic relevance of ACTA2 beyond PeCa, we analyzed public pan-squamous carcinoma cohorts. Higher ACTA2 expression was associated with poorer survival, suggesting a conserved association between myCAF-related signatures and adverse outcomes across squamous carcinoma types ($p = 0.0066$, Figure 5G). Patients with lower ACTA2 expression showed markedly better survival (HR = 0.78, 95% CI: 0.65-0.93, $p = 0.007$, Figure 5H), supporting the potential relevance of myCAF-associated signatures in squamous carcinoma progression.

Cell-cell communication analysis characterized predicted interactions between CAF subtypes and other microenvironmental components. S07_myCAF_ACTA2 primarily engaged with vascular endothelial and epithelial cells through VEGFA-VEGFR, TGFB-TGFBR, ANGPT1-TEK and PLAU-PLAUR signaling axes. Notably, CellChat analysis in patients with early lymph node metastasis predicted POSTN-(ITGAV+ITGB5) signaling between S07_myCAF_ACTA2 and malignant epithelial cells, suggesting a potential link to EMT (Figure 5I). Validation of key communication molecules revealed distinct expression patterns of POSTN, ITGAV and ITGB5 across CAF subtypes, with POSTN showing 95.1% positivity in S07_myCAF_ACTA2, compared to 45.0% and 25.5% for ITGAV and ITGB5, respectively (Figure 5J). Spatial expression maps showed predominantly stromal POSTN expression and more widespread ITGAV and ITGB5 expression (Figure 5K). However, in the analyzed lymph-node section, POSTN-high spots were not significantly closer to ITGAV- or ITGB5-high spots than expected from 999 random spot sets (observed/random nearest-neighbor distance ratios of 1.41 and 1.26, respectively; both P = 1.0; Figure S4j). Thus, these spatial data describe tissue expression patterns but do not provide evidence for preferential nearest-neighbor proximity between POSTN-high and receptor-high spots under the tested conditions.

BHLHE41-mediated POSTN expression in myCAFs is associated with lymph node metastatic features in penile cancer.

Having characterized the functional properties of S07_myCAF_ACTA2, we further investigated its transcriptional regulatory network. TF enrichment analysis identified several regulators enriched in S07_myCAF_ACTA2, with BHLHE41 showing the strongest upregulation (Figure 6A). As a hypoxia-inducible TF, BHLHE41 has been established to modulate cell proliferation, apoptosis and metastasis across various cancers [35]. Other pivotal regulators included stress-response modulator ATF4, inflammatory master regulator NFKB1, prostate development factor NKX3-1, and embryonic regulator SOX4, whose coordinated actions likely drove the pro-metastatic phenotype of the S07_myCAF_ACTA2 subtype.

Expression profiling confirmed elevated BHLHE41 levels in S07_myCAF_ACTA2 compared to S06_iCAF_IL6 and S08_apCAF_CD74 (Figure 6B). Spatial analysis of the representative case showed predominant BHLHE41 expression in peritumoral stromal regions, overlapping spatially with the distribution of S07_myCAF_ACTA2 (Figure 6C). Correlation analysis showed a significant positive association between BHLHE41 and POSTN ($R = 0.22$, $p < 0.001$) (Figure 6D), suggesting that BHLHE41 may contribute to stromal remodeling through regulation of POSTN expression. Previous studies have shown that the extracellular-matrix glycoprotein POSTN can

promote CAF activation, collagen deposition, and tumor invasion in several tumor contexts [36-38]. POSTN binds to integrin receptors ITGAV/ITGB5 to activate the PI3K/AKT and Wnt signaling pathways, thereby promoting EMT and metastasis of tumor cells. Our findings implicated the BHLHE41-POSTN axis as a potential regulatory mechanism contributing to the pro-invasive effects of S07_myCAF_ACTA2 observed in vitro.

Exploratory survival analysis in pan-squamous carcinoma cohorts revealed that elevated expression of POSTN and BHLHE41 was associated with poorer prognosis. Patients with high POSTN expression had significantly shorter survival ($p = 0.0016$), while higher BHLHE41 expression was also associated with adverse outcomes ($p = 0.044$) (Figure 6E). Pan-cancer analysis demonstrated consistent upregulation of BHLHE41 and POSTN across multiple SCC types, including nodal, cervical, distant metastatic, head and neck, laryngeal, oral and oropharyngeal tumors (Figure 6G), suggesting that these factors may represent conserved stromal-associated features across squamous carcinomas. These findings suggest that BHLHE41 may participate in the regulation of POSTN expression in S07_myCAF_ACTA2, potentially contributing to a pro-invasive stromal state through downstream targets such as POSTN.

BHLHE41 in penile cancer CAFs may regulate the epithelial-mesenchymal transition of tumor cells by modulating POSTN and subsequently acting on ITGAV and ITGB5 in tumor cells.

Primary PeCa cells and cancer-associated fibroblasts (CAFs) were isolated from samples collected at our center and characterized as described in the Methods. Representative morphologies of the two cell populations are shown in Figure 7A. Immunofluorescence staining of cultured CAFs showed ACTA2 and COL1A1 expression (Figure 7B). Multiplex immunofluorescence staining showed the spatial distribution of BHLHE41, POSTN, ACTA2, ITGAV, and ITGB5 in PeCa tissues (Figure 7C). POSTN staining was observed in regions containing ACTA2-positive myofibroblastic CAFs, whereas ITGAV and ITGB5 staining was observed in tumor-cell regions.

To investigate the functional relevance of the BHLHE41-POSTN-(ITGAV+ITGB5) signaling axis, multiple co-culture systems were established. In Transwell assays, PeCa cells co-cultured with BHLHE41-overexpressing CAFs showed greater invasive capacity than cells co-cultured with BHLHE41-knockdown CAFs (Figure 8A, S4A-B). After co-culture, PeCa cells were separately harvested and used for subsequent analyses. For Western blot analysis, the harvested PeCa cells were lysed and centrifuged, and the resulting cell lysates were used to assess the expression of the EMT markers N-cadherin, Vimentin, and E-cadherin. BHLHE41 overexpression increased N-cadherin and Vimentin protein levels and decreased E-cadherin protein levels (Figure 8B-E). These findings support an association between BHLHE41 expression in CAFs and EMT-related changes in co-cultured PeCa cells. After co-culture, PeCa cells were also collected and subjected to total RNA extraction, followed by qPCR analysis. The qPCR results further showed consistent changes in the transcriptional levels of N-cadherin, Vimentin, and E-cadherin (Figure 8F-H). To further examine the relationship between BHLHE41 and POSTN in CAFs, POSTN expression was evaluated in the CAF population after co-culture. CAFs were separately harvested for total RNA extraction and subsequent qPCR analysis to assess POSTN transcriptional levels. In parallel, separately harvested CAFs were lysed and centrifuged, and the resulting cell lysates were used for Western blot analysis to quantify POSTN protein expression. Both POSTN mRNA and protein levels in CAFs were significantly increased under BHLHE41-overexpression conditions (Figure 8I-K).

Finally, to evaluate the functional involvement of POSTN-related integrin signaling in EMT regulation, four co-culture systems were established: PeCa cells alone, PeCa cells co-cultured with CAFs, PeCa cells co-cultured with CAFs in the presence of the integrin-blocking Arg-Gly-Asp (RGD) peptide, and PeCa cells co-cultured with POSTN-knockdown CAFs. After co-culture, PeCa cells were separately harvested from each experimental condition for RNA extraction and qPCR analysis, while parallel PeCa-cell samples were lysed for Western blot analysis. Compared with PeCa cells co-cultured with untreated CAFs, both POSTN knockdown in CAFs and RGD treatment resulted in similar changes in EMT marker expression in PeCa cells. Western blot analysis showed that N-cadherin and Vimentin protein levels decreased, whereas E-cadherin protein levels increased (Figure 8O-R). Consistently, qPCR analysis showed decreased N-cadherin and Vimentin mRNA levels and increased E-cadherin mRNA levels under POSTN knockdown or RGD treatment conditions (Figure 8L-N).

Discussion

PeCa is a rare but aggressive malignancy of the male genitourinary tract, and lymph node metastasis is the most important prognostic factor [2,39]. Because PeCa is rare, its molecular mechanisms of metastasis remain incompletely understood, particularly the contribution of the tumor microenvironment. In this study, we systematically analyzed the cellular heterogeneity of the PeCa tumor microenvironment and its role in early lymph node metastasis for the first time by integrating single-cell RNA sequencing and spatial transcriptomics.

Our study identified cellular changes associated with metastatic PeCa. Among malignant epithelial cells, metastasis-associated malignant cell states showed higher stemness, enrichment of nervous-system-development programs, and reduced MHC-related signaling. These malignant cell states also showed strong hypoxia-associated signatures, reflecting a synergistic pattern of tumor cell dedifferentiation, metabolic reprogramming and immune evasion. In terms of the immune microenvironment, the proportions of B cells and CD4 T cells were reduced in metastatic tissues, whereas immunosuppressive Treg cells and T-cell subsets with high expression of interferon-stimulated genes were enriched, consistent with chronic antigen stimulation and functional exhaustion. These changes collectively form a coordinated network characterized by enhanced cancer cell stemness and exacerbated immunosuppression, creating favorable conditions for tumor metastasis.

We further identified S07_myCAF_ACTA2 as a CAF subset associated with the pro-metastatic microenvironment in PeCa. This subset was enriched in metastatic tissues and exhibited prominent hypoxia-associated features. It was enriched in niche_9, which had a relatively high hypoxia score, and was observed in both the primary and metastatic lesions of the spatially profiled case. At the molecular level, we found that BHLHE41, a hypoxia-inducible transcription factor, showed preferential activity in S07_myCAF_ACTA2 and may contribute to the regulation of POSTN expression. Notably, the biological role of BHLHE41 appears to be context-dependent across different tumor types. Previous studies have reported tumor-suppressive effects of BHLHE41/DEC2/SHARP1 in pancreatic and breast cancers [40,41], suggesting that its functional role may vary according to tumor type and cellular context. In PeCa, our findings suggest that BHLHE41 activation in myCAF_ACTA2 is associated with POSTN-related pro-metastatic stromal signaling. Previous studies provide evidence that stromal POSTN can function as a mediator of tumor-CAF crosstalk in other malignancies. In ovarian cancer, TGF-β1-induced POSTN derived

from activated fibroblasts promoted tumor-cell migration and invasion through integrin-mediated PI3K/AKT signaling [42]. Importantly, a similar stromal role for POSTN has been reported in head and neck cancer, where CAF-derived POSTN was increased in tumors with lymph node metastasis and promoted cancer-cell growth, migration, and invasion. These findings provide a general mechanistic precedent for POSTN-mediated tumor-stromal crosstalk [38]. Spatial transcriptomic analysis of the representative metastatic PeCa case suggested peritumoral enrichment of S07_myCAF_ACTA2. However, the nearest-neighbor analysis did not detect preferential proximity of POSTN-high spots to ITGAV- or ITGB5-high spots in the analyzed lymph-node section (Figure S4J). Accordingly, the spatial findings do not establish preferential ligand–receptor proximity at the niche level. Several biological and technical factors may contribute to this lack of detectable nearest-neighbor proximity. POSTN is a secreted extracellular-matrix protein, and its extracellular distribution and site of biological activity may not precisely coincide with the spatial location of POSTN-expressing cells or transcripts. Moreover, incorporation and retention of POSTN within the extracellular matrix may permit paracrine effects beyond immediately adjacent transcript-positive spots. Spatial transcriptomic analysis is also based on a two-dimensional section of a three-dimensional tissue architecture; therefore, cells that are adjacent in three-dimensional space may not necessarily appear as nearest neighbors within a single section. These factors may partly contribute to the observed spatial pattern, but they do not provide evidence for preferential POSTN – ITGAV/ITGB5 proximity. The proposed POSTN-related CAF–tumor signaling model is supported by cell-cell communication predictions and in vitro perturbation experiments, whereas its spatial organization and receptor-specific activity in situ remain to be established. We further assessed the association between BHLHE41 and POSTN using immunofluorescence assays, WB, and qPCR. We then examined the relationship between POSTN-(ITGAV+ITGB5) signaling and epithelial-mesenchymal transition (EMT) in PeCa and evaluated its functional relevance using Transwell assays.

These findings have potential scientific and translational implications. First, we revealed a molecular link between the BHLHE41-POSTN transcriptional regulatory axis and hypoxic microenvironment, CAF activation and tumor metastasis. Second, the identification of metastasis-associated cancer cell states and hypoxia-adapted CAF characteristics provides potential insights for precise molecular classification of PeCa and development of microenvironment-based therapeutic strategies. From a clinical perspective, S07_myCAF_ACTA2 and its associated niche characteristics may provide potential biomarkers for studying lymph node metastasis risk in PeCa, while the BHLHE41-POSTN axis represents a potential mechanism for future CAF-targeted therapeutic investigation.

This study also suggests molecular commonalities between PeCa and other SCCs, providing a rationale for evaluating the translation of therapeutic strategies across tumor types. However, several limitations should be noted. Although our integrated scRNA-seq cohort included 24 treatment-naïve PeCa patients, the rarity of PeCa limited the cohort size. Moreover, spatial transcriptomic profiling was performed in only one representative metastatic case; therefore, the identified spatial patterns should be considered hypothesis-generating and require validation in larger cohorts. Furthermore, our study lacked in vivo validation and relied mainly on transcriptomic analyses. Although the clinical cohort analyses demonstrated associations between S07_myCAF_ACTA2, the BHLHE41-POSTN axis, and lymph node metastatic status, and our in vitro experiments showed that modulation of this pathway affected EMT and tumor-cell invasion, these findings do not establish

a causal role of this mechanism in lymph node metastasis in vivo. The RGD peptide blockade experiment also has limited specificity for confirming the POSTN-ITGAV/ITGB5 interaction. Future studies should focus on verifying the functional role of the BHLHE41-POSTN axis, further confirming the specific involvement of ITGAV/ITGB5 through targeted knockdown or blockade experiments, evaluating therapeutic strategies based on CAF characteristics, and validating the clinical utility of predictive models in larger cohorts.

Conclusion

This study characterized cellular heterogeneity in the PeCa tumor microenvironment and identified stem-like malignant cell states, immune remodeling, and a hypoxia-adapted myCAF population associated with lymph node metastasis. Our findings further support an in vitro mechanistic model in which the BHLHE41-POSTN axis may contribute to EMT-associated invasive phenotypes in PeCa cells. These results provide a crucial molecular basis and practical tools for the precise diagnosis and treatment of PeCa.

Abbreviations

PeCa	Penile cancer
SCC	Squamous cell carcinoma
CAFs	Cancer-associated fibroblasts
EMT	Epithelial-mesenchymal transition
PSCC	Penile squamous cell carcinoma
CNV	Copy number variation
FFPE	Formalin Fixed Paraffin Embedded
UMIs	Unique molecular identifiers
BANKSY	Building Aggregates with a Neighborhood Kernel and Spatial Yardstick
RCTD	Robust Cell Type Decomposition
Ro/e	Ratio of observed to expected cell number
OR	Odds ratio
TME	Tumor microenvironment
TF	Transcription factor
SCENIC	Single-cell regulatory network inference and clustering
myCAF	Myofibroblastic CAF
UMAP	Uniform Manifold Approximation and Projection
DEGs	Differentially expressed genes
KEGG	Kyoto Encyclopedia of Genes and Genomes

Authors contributions

Conceptualization: Feiya Yang, Haoyuan Shi, Hui Xu, Mo Zhang and Nianzeng Xing.
Methodology: Feiya Yang, Liyuan Wu and Yong Liu.
Software: Haoyuan Shi, Haoxi Wang and Yong Liu.
Validation: Feiya Yang and Yiwei Liu.
Formal analysis: Feiya Yang, Haoyuan Shi, Haoxi Wang and Liyuan Wu.
Investigation: Haoyuan Shi, Yiwei Liu, Huaqi Yin and Dong Chen.
Resources: Mo Zhang, Lu Yin, Huaqi Yin and Kehao Pan.

Data curation: Feiya Yang, Haoyuan Shi, Haoxi Wang and Liyuan Wu.
Writing - original draft: Feiya Yang, Haoyuan Shi, Liyuan Wu, Haoxi Wang and Yong Liu.
Writing - review & editing: Xigao Liu, Tiejun Yang, Hui Xu, Mo Zhang and Nianzeng Xing.
Visualization: Haoxi Wang, Wenkuan Wang and Xigao Liu.
Supervision: Nianzeng Xing.
Project administration: Xigao Liu, Tiejun Yang, Hui Xu, Mo Zhang and Nianzeng Xing.
Funding acquisition: Nianzeng Xing.

Acknowledgments

The authors thank Shanghai Fengshengu Biotechnology Co., Ltd., OE Biotech Co., Ltd. (Shanghai, China), for their assistance in this study.

Funding

This work was supported by National Key R&D Program of China (2022YFE0200800), National Key R&D Program of China (2023YFC2507000), the CAMS Innovation Fund for Medical Sciences (2025-I2M-XHCL-045), Individual Human Project of foreign experts (Category H) (H20240740), the Natural Science Foundation of Beijing,China (L2608059), the National Natural Science Foundation of China (82674027) and the Postdoctoral Fellowship Program of China Postdoctoral Science Foundation (CPSF) (GZC20240147).

Authors disclosures

There is nothing to disclose.

Conflict of interest

All authors declare no conflicts of interest.

Data and code availability

All datasets, data processing codes, and configuration files are stored on GitHub (https://github.com/yuansh3354/scPECA). All scRNA-seq resources in this study are available in the China National Center for Bioinformation (BioProject ID PRJCA044556, accession number HRA019327).

Ethics declarations

Ethics approval and consent to participate

This study was reviewed and approved by the National Cancer Center and the Independent Ethics Committee (Ethics code: NCC2024C-663, NCC2025C-1363).

Consent for publication

All authors have read and approved the publication of this manuscript.

Competing interests

The authors declare no competing interests.

Take Home Message

Hypoxia-inducible BHLHE41 upregulates POSTN in myCAF_ACTA2 and promotes epithelial-mesenchymal transition through ITGAV/ITGB5-related signaling in cancer cells, potentially

contributing to the lymph node metastatic phenotype of PeCa. High BHLHE41 and POSTN levels correlate with poor survival in squamous cell carcinoma patients.

Advancing practice text

What does the study add

Our study provides a single-cell and spatial transcriptomic atlas of penile cancer (PeCa), revealing the heterogeneity of cancer-associated fibroblasts (CAFs) and identifying stromal features associated with lymph node metastasis. The BHLHE41-POSTN axis may have utility as a biomarker framework for metastasis-risk assessment and provides a rationale for future investigation of CAF-directed therapeutic strategies.

Patient Summary

In this study, we analyzed penile carcinoma data from a Chinese cohort and observed that patients with higher myCAF infiltration were more likely to present with early lymph node metastasis.

References

1. Thomas A, Necchi A, Muneer A, Tobias-Machado M, Tran ATH, Van Rompuy AS, et al. Penile cancer. Nat Rev Dis Primers. 2021; 7(1): 11.
2. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024; 74(3): 229-63.
3. Ahmed ME, Khalil MI, Kamel MH, Karnes RJ, Spiess PE. Progress on Management of Penile Cancer in 2020. Curr Treat Options Oncol. 2020; 22(1): 4.
4. Wood HM, Angermeier KW. Anatomic considerations of the penis, lymphatic drainage, and biopsy of the sentinel node. Urol Clin North Am. 2010; 37(3): 327-34.
5. Zhu Y, Ye DW. Lymph node metastases and prognosis in penile cancer. Chin J Cancer Res. 2012; 24(2): 90-6.
6. Ficarra V, Akduman B, Bouchot O, Palou J, Tobias-Machado M. Prognostic factors in penile cancer. Urology. 2010; 76(2 Suppl 1): S66-73.
7. Elst L, Vreeburg M, Brouwer O, Albersen M. Challenges in Organ-sparing Surgery for Penile Cancer: What Are the Limits? Eur Urol Focus. 2023; 9(2): 241-3.
8. Djajadiningrat RS, van Werkhoven E, Horenblas S. Prophylactic pelvic lymph node dissection in patients with penile cancer. J Urol. 2015; 193(6): 1976-80.
9. Deng X, Liu Y, Zhan X, Chen T, Jiang M, Jiang X, et al. Trends in Incidence, Mortality, and Survival of Penile Cancer in the United States: A Population-Based Study. Front Oncol. 2022; 12: 891623.
10. Skeppner E, Andersson SO, Johansson JE, Windahl T. Initial symptoms and delay in patients with penile carcinoma. Scand J Urol Nephrol. 2012; 46(5): 319-25.
11. Fu L, Tian T, Yao K, Chen XF, Luo G, Gao Y, et al. Global Pattern and Trends in Penile Cancer Incidence: Population-Based Study. JMIR Public Health Surveill. 2022; 8(7): e34874.
12. Brown A, Ma Y, Danenberg K, Schuckman AK, Pinski JK, Pagliaro LC, et al. Epidermal growth factor receptor-targeted therapy in squamous cell carcinoma of the penis: a report of 3 cases. Urology. 2014; 83(1): 159-65.
13. Carthon BC, Ng CS, Pettaway CA, Pagliaro LC. Epidermal growth factor receptor-targeted therapy in locally advanced or metastatic squamous cell carcinoma of the penis. BJU Int. 2014; 113(6): 871-7.
14. Necchi A, Lo Vullo S, Perrone F, Raggi D, Giannatempo P, Calareso G, et al. First-line therapy with dacomitinib, an orally available pan-HER tyrosine kinase inhibitor, for locally advanced or metastatic penile squamous cell carcinoma: results of an open-label, single-arm, single-centre, phase 2 study. BJU Int. 2018; 121(3): 348-56.
15. Canter DJ, Nicholson S, Watkin N, Hall E, Pettaway C. The International Penile Advanced Cancer Trial (InPACT): Rationale and Current Status. Eur Urol Focus. 2019; 5(5): 706-9.
16. Robinson R, Marconi L, MacPepple E, Hakenberg OW, Watkin N, Yuan Y, et al. Risks and Benefits of Adjuvant Radiotherapy After Inguinal Lymphadenectomy in Node-positive Penile Cancer: A Systematic Review by the European Association of Urology Penile Cancer Guidelines Panel. Eur Urol. 2018; 74(1): 76-83.
17. Zhu Y, Ye DW, Yao XD, Zhang SL, Dai B, Zhang HL, et al. The value of squamous cell carcinoma antigen in the prognostic evaluation, treatment monitoring and followup of patients with penile cancer. J Urol. 2008; 180(5): 2019-23.

18. Zhu Y, Zhou XY, Yao XD, Dai B, Ye DW. The prognostic significance of p53, Ki-67, epithelial cadherin and matrix metalloproteinase-9 in penile squamous cell carcinoma treated with surgery. BJU Int. 2007; 100(1): 204-8.
19. Tan X, Liu Z, Wang Y, Wu Z, Zou Y, Luo S, et al. miR-138-5p-mediated HOXD11 promotes cell invasion and metastasis by activating the FN1/MMP2/MMP9 pathway and predicts poor prognosis in penile squamous cell carcinoma. Cell Death Dis. 2022; 13(9): 816.
20. Takahashi K, Podyma-Inoue KA, Saito M, Sakakitani S, Sugauchi A, Iida K, et al. TGF-β generates a population of cancer cells residing in G1 phase with high motility and metastatic potential via KRTAP2-3. Cell Rep. 2022; 40(13): 111411.
21. Bordignon P, Bottoni G, Xu X, Popescu AS, Truan Z, Guenova E, et al. Dualism of FGF and TGF-β Signaling in Heterogeneous Cancer-Associated Fibroblast Activation with ETV1 as a Critical Determinant. Cell Rep. 2019; 28(9): 2358-72.e6.
22. Liu Y, Sinjab A, Min J, Han G, Paradiso F, Zhang Y, et al. Conserved spatial subtypes and cellular neighborhoods of cancer-associated fibroblasts revealed by single-cell spatial multi-omics. Cancer Cell. 2025; 43(5): 905-24.e6.
23. Wang Y, Yang W, Wang Q, Zhou Y. Mechanisms of esophageal cancer metastasis and treatment progress. Front Immunol. 2023; 14: 1206504.
24. Chen J, Larsson L, Swarbrick A, Lundeberg J. Spatial landscapes of cancers: insights and opportunities. Nat Rev Clin Oncol. 2024; 21(9): 660-74.
25. Elst L, Philips G, Vandermaesen K, Bassez A, Lodi F, Vreeburg MTA, et al. Single-cell Atlas of Penile Cancer Reveals TP53 Mutations as a Driver of an Aggressive Phenotype, Irrespective of Human Papillomavirus Status, and Provides Clues for Treatment Personalization. Eur Urol. 2024; 86(2): 114-27.
26. Singhal V, Chou N, Lee J, Yue Y, Liu J, Chock WK, et al. BANKSY unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis. Nat Genet. 2024; 56(3): 431-41.
27. Corvino D, Batstone M, Hughes BGM, Kempchen T, Ng SS, Salim N, et al. Type I Interferon Drives a Cellular State Inert to TCR-Stimulation and Could Impede Effective T-Cell Differentiation in Cancer. Eur J Immunol. 2025; 55(1): e202451371.
28. Acha-Sagredo A, Andrei P, Clayton K, Taggart E, Antoniotti C, Woodman CA, et al. A constitutive interferon-high immunophenotype defines response to immunotherapy in colorectal cancer. Cancer Cell. 2025; 43(2): 292-307.e7.
29. Calvo F, Ege N, Grande-Garcia A, Hooper S, Jenkins RP, Chaudhry SI, et al. Mechanotransduction and YAP-dependent matrix remodelling is required for the generation and maintenance of cancer-associated fibroblasts. Nat Cell Biol. 2013; 15(6): 637-46.
30. Chen X, Song E. The theory of tumor ecosystem. Cancer Commun (Lond). 2022; 42(7): 587-608.
31. Payne LB, Hoque M, Houk C, Darden J, Chappell JC. Pericytes in Vascular Development. Curr Tissue Microenviron Rep. 2020; 1(3): 143-54.
32. Barnett SN, Cujba AM, Yang L, Maceiras AR, Li S, Kedlian VR, et al. An organotypic atlas of human vascular cells. Nat Med. 2024; 30(12): 3468-81.
33. Crawford KS, Volkman BF. Prospects for targeting ACKR1 in cancer and other diseases. Front Immunol. 2023; 14: 1111960.
34. Mucciolo G, Araos Henríquez J, Jihad M, Pinto Teles S, Manansala JS, Li W, et al. EGFR-

activated myofibroblasts promote metastasis of pancreatic cancer. Cancer Cell. 2024; 42(1): 101-18.e11.
35. Bret C, Desmots-Loyer F, Moreaux J, Fest T. BHLHE41, a transcriptional repressor involved in physiological processes and tumor development. Cell Oncol (Dordr). 2025; 48(1): 43-66.
36. Null JL, Kim DJ, McCann JV, Pramoonjago P, Fox JW, Zeng J, et al. Periostin+ Stromal Cells Guide Lymphovascular Invasion by Cancer Cells. Cancer Res. 2023; 83(13): 2105-22.
37. Neuzillet C, Nicolle R, Raffenne J, Tijeras-Raballand A, Brunel A, Astorgues-Xerri L, et al. Periostin- and podoplanin-positive cancer-associated fibroblast subtypes cooperate to shape the inflamed tumor microenvironment in aggressive pancreatic adenocarcinoma. J Pathol. 2022; 258(4): 408-25.
38. Qin X, Yan M, Zhang J, Wang X, Shen Z, Lv Z, et al. TGFβ3-mediated induction of Periostin facilitates head and neck cancer growth and is associated with metastasis. Sci Rep. 2016; 6: 20587.
39. Brouwer OR, Albersen M, Parnham A, Protzel C, Pettaway CA, Ayres B, et al. European Association of Urology-American Society of Clinical Oncology Collaborative Guideline on Penile Cancer: 2023 Update. Eur Urol. 2023; 83(6): 548-60.
40. Sato F, Kawamura H, Wu Y, Sato H, Jin D, Bhawal UK, et al. The basic helix-loop-helix transcription factor DEC2 inhibits TGF-β-induced tumor progression in human pancreatic cancer BxPC-3 cells. Int J Mol Med. 2012; 30(3): 495-501.
41. Montagner M, Enzo E, Forcato M, Zanconato F, Parenti A, Rampazzo E, et al. SHARP1 suppresses breast cancer metastasis by promoting degradation of hypoxia-inducible factors. Nature. 2012; 487(7407): 380-4.
42. Yue H, Li W, Chen R, Wang J, Lu X, Li J. Stromal POSTN induced by TGF-β1 facilitates the migration and invasion of ovarian cancer. Gynecol Oncol. 2021; 160(2): 530-8.

Figure 1

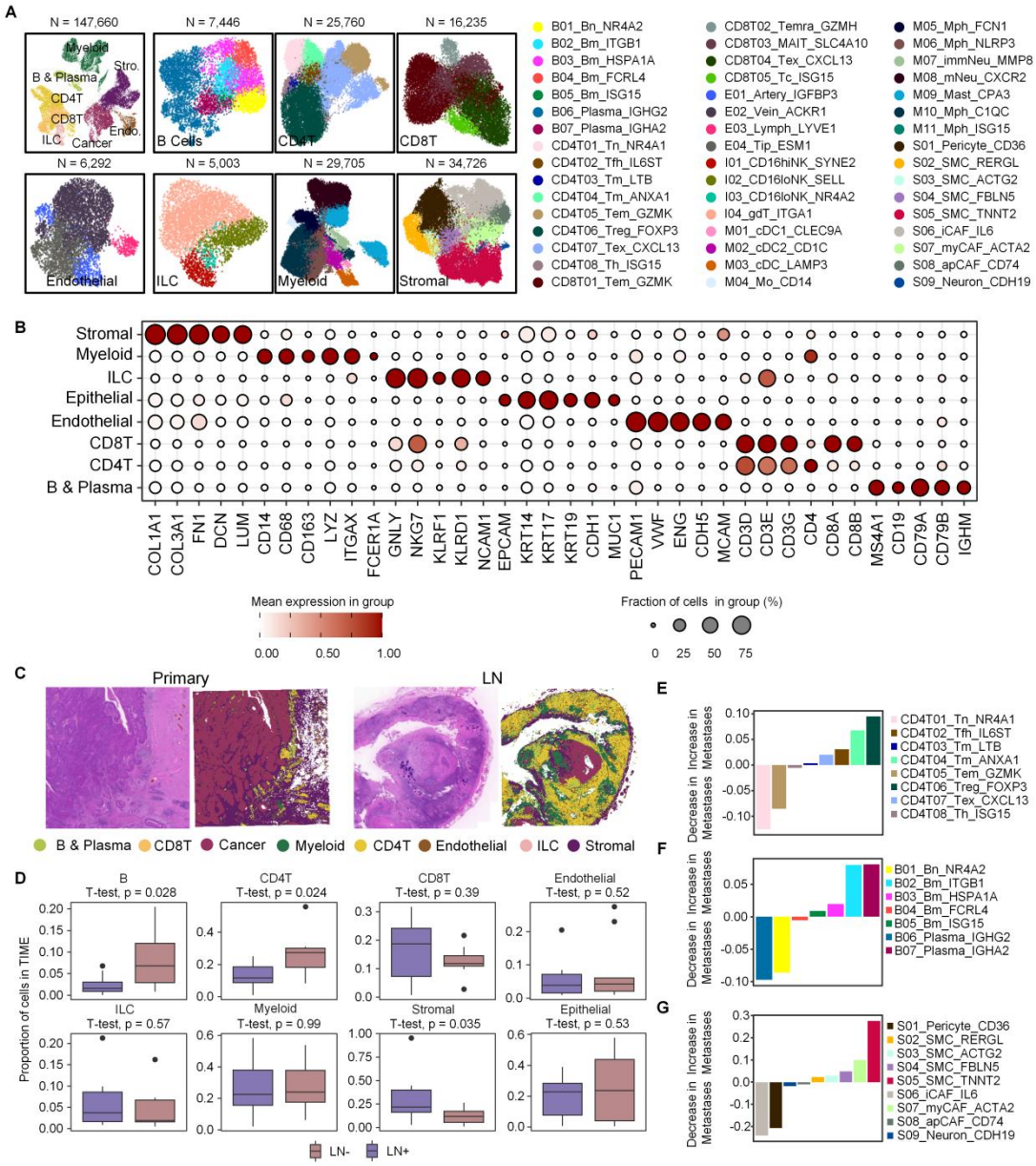

A
N = 147,660
N = 7,446
N = 25,760
N = 16,235
Myeloid
Stro.
B & Plasma
CD4T
CD8T
ILC
Cancer
Endo.
B Cells
CD4T
CD8T
N = 6,292
N = 5,003
N = 29,705
N = 34,726
Endothelial
ILC
Myeloid
Stromal
B01_Bn_NR4A2
B02_Bm_ITGB1
B03_Bm_HSPA1A
B04_Bm_FCRL4
B05_Bm_ISG15
B06_Plasma_IGHG2
B07_Plasma_IGHA2
CD4T01_Tn_NR4A1
CD4T02_Tfh_IL6ST
CD4T03_Tm_LTB
CD4T04_Tm_ANXA1
CD4T05_Tem_GZMK
CD4T06_Treg_FOXP3
CD4T07_Tex_CXCL13
CD4T08_Th_ISG15
CD8T01_Tem_GZMK
CD8T02_Temra_GZMH
CD8T03_MAIT_SLC4A10
CD8T04_Tex_CXCL13
CD8T05_Tc_ISG15
E01_Artery_IGFBP3
E02_Vein_ACKR1
E03_Lymph_LYVE1
E04_Tip_ESM1
I01_CD16hiNK_SYNE2
I02_CD16loNK_SELL
I03_CD16loNK_NR4A2
I04_gdT_ITGA1
M01_cDC1_CLEC9A
M02_cDC2_CD1C
M03_cDC_LAMP3
M04_Mo_CD14
M05_Mph_FCN1
M06_Mph_NLRP3
M07_immNeu_MMP8
M08_mNeu_CXCR2
M09_Mast_CPA3
M10_Mph_C1QC
M11_Mph_ISG15
S01_Pericyte_CD36
S02_SMC_RERGL
S03_SMC_ACTG2
S04_SMC_FBLN5
S05_SMC_TNNT2
S06_iCAF_IL6
S07_myCAF_ACTA2
S08_apCAF_CD74
S09_Neuron_CDH19
B
Stromal
Myeloid
ILC
Epithelial
Endothelial
CD8T
CD4T
B & Plasma
COL1A1
COL3A1
FN1
DCN
LUM
CD14
CD68
CD163
LYZ
ITGAX
FCER1A
GNLY
NKG7
KLRF1
KLRD1
NCAM1
EPCAM
KRT14
KRT17
KRT19
CDH1
MUC1
PECAM1
VWF
ENG
CDH5
MCAM
CD3D
CD3E
CD3G
CD4
CD8A
CD8B
MS4A1
CD19
CD79A
CD79B
IGHM
Mean expression in group
0.00
0.50
1.00
Fraction of cells in group (%)
0
25
50
75
C
Primary
LN
B & Plasma
CD8T
Cancer
Myeloid
CD4T
Endothelial
ILC
Stromal
D
B
T-test, p = 0.028
CD4T
T-test, p = 0.024
CD8T
T-test, p = 0.39
Endothelial
T-test, p = 0.52
ILC
T-test, p = 0.57
Myeloid
T-test, p = 0.99
Stromal
T-test, p = 0.035
Epithelial
T-test, p = 0.53
Proportion of cells in TIME
LN-
LN+
E
Decrease in Metastases
Increase in Metastases
CD4T01_Tn_NR4A1
CD4T02_Tfh_IL6ST
CD4T03_Tm_LTB
CD4T04_Tm_ANXA1
CD4T05_Tem_GZMK
CD4T06_Treg_FOXP3
CD4T07_Tex_CXCL13
CD4T08_Th_ISG15
F
Decrease in Metastases
Increase in Metastases
B01_Bn_NR4A2
B02_Bm_ITGB1
B03_Bm_HSPA1A
B04_Bm_FCRL4
B05_Bm_ISG15
B06_Plasma_IGHG2
B07_Plasma_IGHA2
G
Decrease in Metastases
Increase in Metastases
S01_Pericyte_CD36
S02_SMC_RERGL
S03_SMC_ACTG2
S04_SMC_FBLN5
S05_SMC_TNNT2
S06_iCAF_IL6
S07_myCAF_ACTA2
S08_apCAF_CD74
S09_Neuron_CDH19

Figure 2

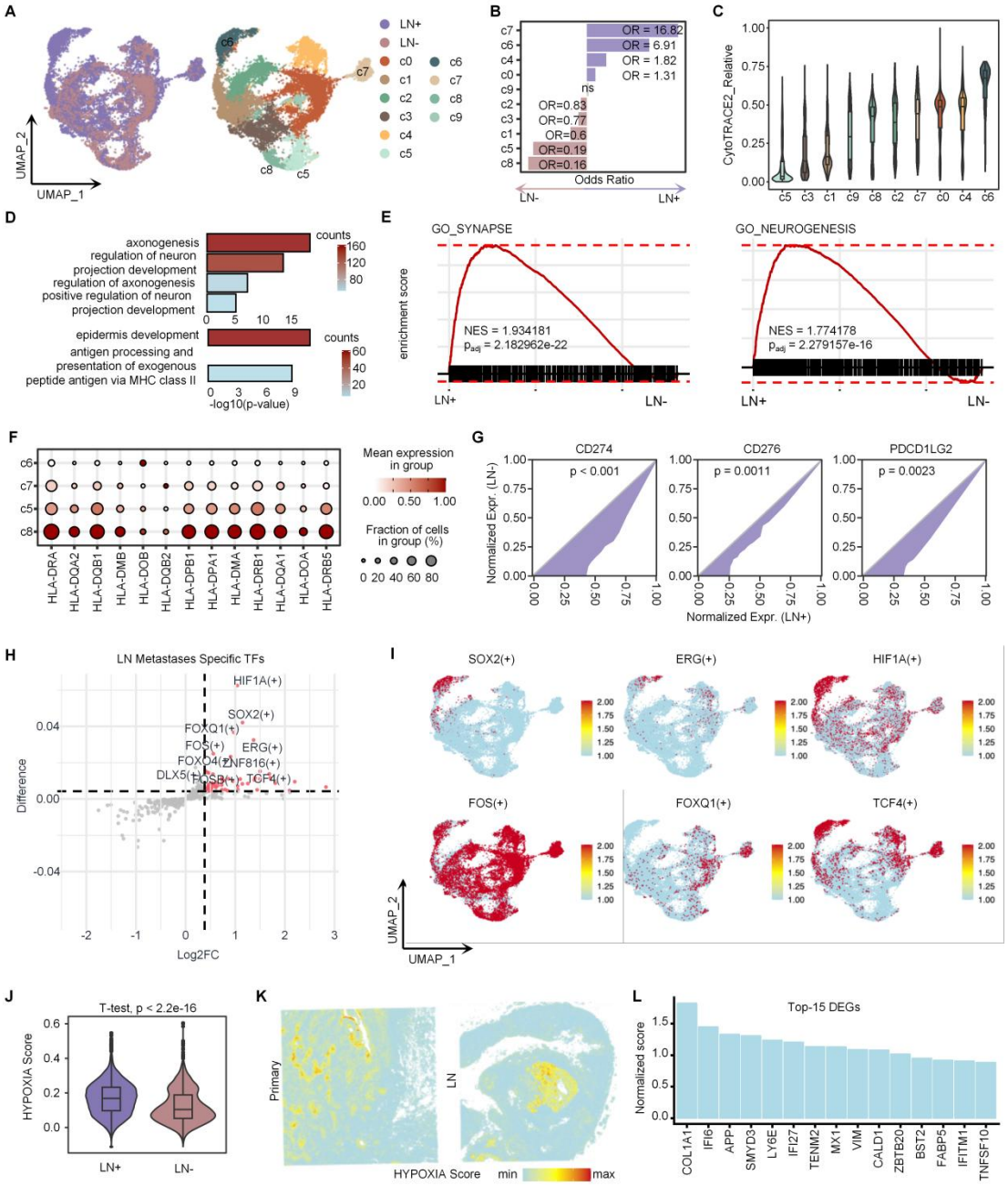

A
LN+
LN-
c0
c1
c2
c3
c4
c5
c6
c7
c8
c9
UMAP_1
UMAP_2
B
OR = 16.82
OR = 6.91
OR = 1.82
OR = 1.31
ns
OR=0.83
OR=0.77
OR=0.6
OR=0.19
OR=0.16
Odds Ratio
C
CytoTRACE2_Relative
D
axonogenesis
regulation of neuron projection development
regulation of axonogenesis
positive regulation of neuron projection development
epidermis development
antigen processing and presentation of exogenous peptide antigen via MHC class II
counts
-log10(p-value)
E
GO_SYNAPSE
GO_NEUROGENESIS
enrichment score
NES = 1.934181
padj = 2.182962e-22
NES = 1.774178
padj = 2.279157e-16
F
Mean expression in group
Fraction of cells in group (%)
HLA-DRA
HLA-DQA2
HLA-DQB1
HLA-DMB
HLA-DOB
HLA-DQB2
HLA-DPB1
HLA-DPA1
HLA-DMA
HLA-DRB1
HLA-DQA1
HLA-DOA
HLA-DRB5
G
CD274
CD276
PDCD1LG2
p < 0.001
p = 0.0011
p = 0.0023
Normalized Expr. (LN-)
Normalized Expr. (LN+)
H
LN Metastases Specific TFs
HIF1A(+)
SOX2(+)
FOXQ1(+)
FOS(+)
ERG(+)
FOXO4(+)
ZNF816(+)
DLX5(+)
TCF4(+)
Difference
Log2FC
I
SOX2(+)
ERG(+)
HIF1A(+)
FOS(+)
FOXQ1(+)
TCF4(+)
UMAP_1
UMAP_2
J
T-test, p < 2.2e-16
HYPOXIA Score
LN+
LN-
K
Primary
LN
HYPOXIA Score
min
max
L
Top-15 DEGs
Normalized score
COL1A1
IFI6
APP
SMYD3
LY6E
IFI27
TENM2
MX1
VIM
CALD1
ZBTB20
BST2
FABP5
IFITM1
TNFSF10

Figure 3

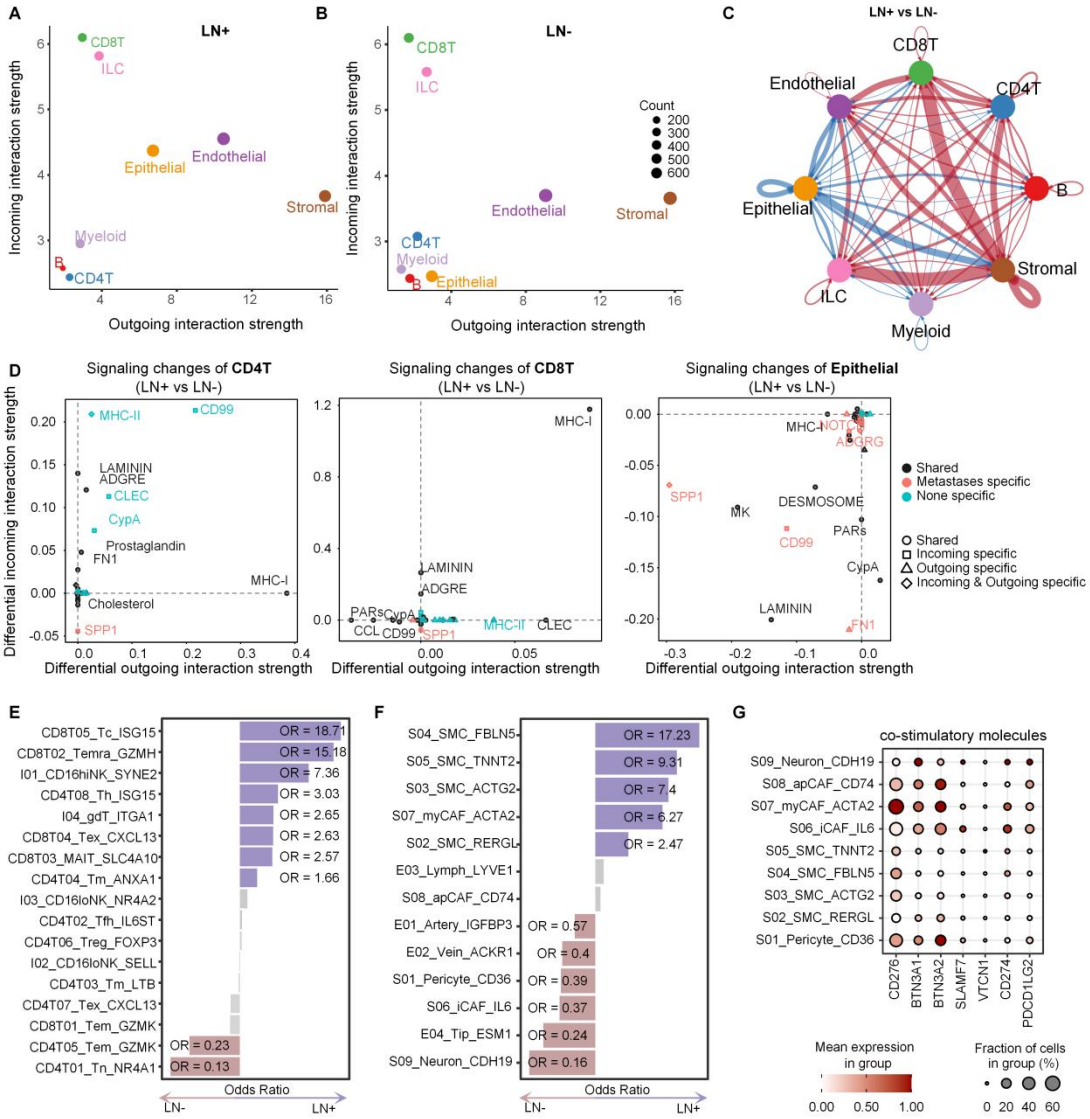

A
LN+
Incoming interaction strength
Outgoing interaction strength
CD8T
ILC
Epithelial
Endothelial
Stromal
Myeloid
B
CD4T
B
LN-
Count
200
300
400
500
600
C
LN+ vs LN-
CD8T
Endothelial
CD4T
Epithelial
B
Stromal
ILC
Myeloid
D
Signaling changes of CD4T (LN+ vs LN-)
Signaling changes of CD8T (LN+ vs LN-)
Signaling changes of Epithelial (LN+ vs LN-)
Differential incoming interaction strength
Differential outgoing interaction strength
MHC-II
CD99
LAMININ
ADGRE
CLEC
CypA
Prostaglandin
FN1
MHC-I
Cholesterol
SPP1
PARs
CCL
MHC-II
NOTCH
ADGRG
DESMOSOME
MK
CD99
Shared
Metastases specific
None specific
Shared
Incoming specific
Outgoing specific
Incoming & Outgoing specific
E
CD8T05_Tc_ISG15
CD8T02_Temra_GZMH
I01_CD16hiNK_SYNE2
CD4T08_Th_ISG15
I04_gdT_ITGA1
CD8T04_Tex_CXCL13
CD8T03_MAIT_SLC4A10
CD4T04_Tm_ANXA1
I03_CD16loNK_NR4A2
CD4T02_Tfh_IL6ST
CD4T06_Treg_FOXP3
I02_CD16loNK_SELL
CD4T03_Tm_LTB
CD4T07_Tex_CXCL13
CD8T01_Tem_GZMK
CD4T05_Tem_GZMK
CD4T01_Tn_NR4A1
OR = 18.71
OR = 15.18
OR = 7.36
OR = 3.03
OR = 2.65
OR = 2.63
OR = 2.57
OR = 1.66
OR = 0.23
OR = 0.13
Odds Ratio
LN-
LN+
F
S04_SMC_FBLN5
S05_SMC_TNNT2
S03_SMC_ACTG2
S07_myCAF_ACTA2
S02_SMC_RERGL
E03_Lymph_LYVE1
S08_apCAF_CD74
E01_Artery_IGFBP3
E02_Vein_ACKR1
S01_Pericyte_CD36
S06_iCAF_IL6
E04_Tip_ESM1
S09_Neuron_CDH19
OR = 17.23
OR = 9.31
OR = 7.4
OR = 6.27
OR = 2.47
OR = 0.57
OR = 0.4
OR = 0.39
OR = 0.37
OR = 0.24
OR = 0.16
Odds Ratio
LN-
LN+
G
co-stimulatory molecules
S09_Neuron_CDH19
S08_apCAF_CD74
S07_myCAF_ACTA2
S06_iCAF_IL6
S05_SMC_TNNT2
S04_SMC_FBLN5
S03_SMC_ACTG2
S02_SMC_RERGL
S01_Pericyte_CD36
CD276
BTN3A1
BTN3A2
SLAMF7
VTCN1
CD274
PDCD1LG2
Mean expression in group
0.00 0.50 1.00
Fraction of cells in group (%)
0 20 40 60

Figure 4

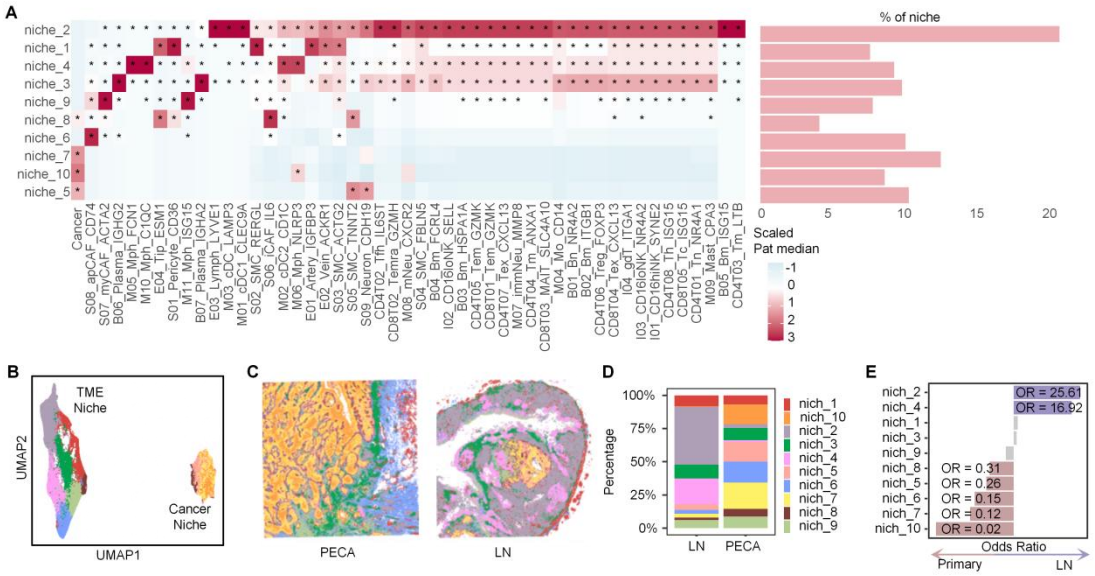

A
niche_2
niche_1
niche_4
niche_3
niche_9
niche_8
niche_6
niche_7
niche_10
niche_5
% of niche
0
5
10
15
20
Scaled
Pat median
-1
0
1
2
3
Cancer
S08_apCAF_CD74
S07_myCAF_ACTA2
B06_Plasma_IGHG2
M05_Mph_FCN1
M10_Mph_C1QC
E04_Tip_ESM1
S01_Pericyte_CD36
M11_Mph_ISG15
B07_Plasma_IGHA2
E03_Lymph_LYVE1
M03_cDC_LAMP3
M01_cDC1_CLEC9A
S02_SMC_RERGL
S06_iCAF_IL6
M02_cDC2_CD1C
M06_Mph_NLRP3
E01_Artery_IGFBP3
E02_Vein_ACKR1
S03_SMC_ACTG2
S05_SMC_TNNT2
S09_Neuron_CDH19
CD4T02_Tfh_IL6ST
CD8T02_Temra_GZMH
M08_mNeu_CXCR2
S04_SMC_FBLN5
B04_Bm_FCRL4
I02_CD16loNK_SELL
B03_Bm_HSPA1A
CD4T05_Tem_GZMK
CD8T01_Tem_GZMK
CD4T07_Tex_CXCL13
M07_immNeu_MMP8
CD4T04_Tm_ANXA1
CD8T03_MAIT_SLC4A10
M04_Mo_CD14
B01_Bn_NR4A2
B02_Bm_ITGB1
CD4T06_Treg_FOXP3
CD8T04_Tex_CXCL13
I04_gdT_ITGA1
I03_CD16loNK_NR4A2
I01_CD16hiNK_SYNE2
CD4T08_Th_ISG15
CD8T05_Tc_ISG15
CD4T01_Tn_NR4A1
M09_Mast_CPA3
B05_Bm_ISG15
CD4T03_Tm_LTB
B
TME
Niche
Cancer
Niche
UMAP2
UMAP1
C
PECA
LN
D
100%
75%
50%
25%
0%
Percentage
LN
PECA
nich_1
nich_10
nich_2
nich_3
nich_4
nich_5
nich_6
nich_7
nich_8
nich_9
E
nich_2
nich_4
nich_1
nich_3
nich_9
nich_8
nich_5
nich_6
nich_7
nich_10
OR = 25.61
OR = 16.92
OR = 0.31
OR = 0.26
OR = 0.15
OR = 0.12
OR = 0.02
Odds Ratio
Primary
LN

Figure 5

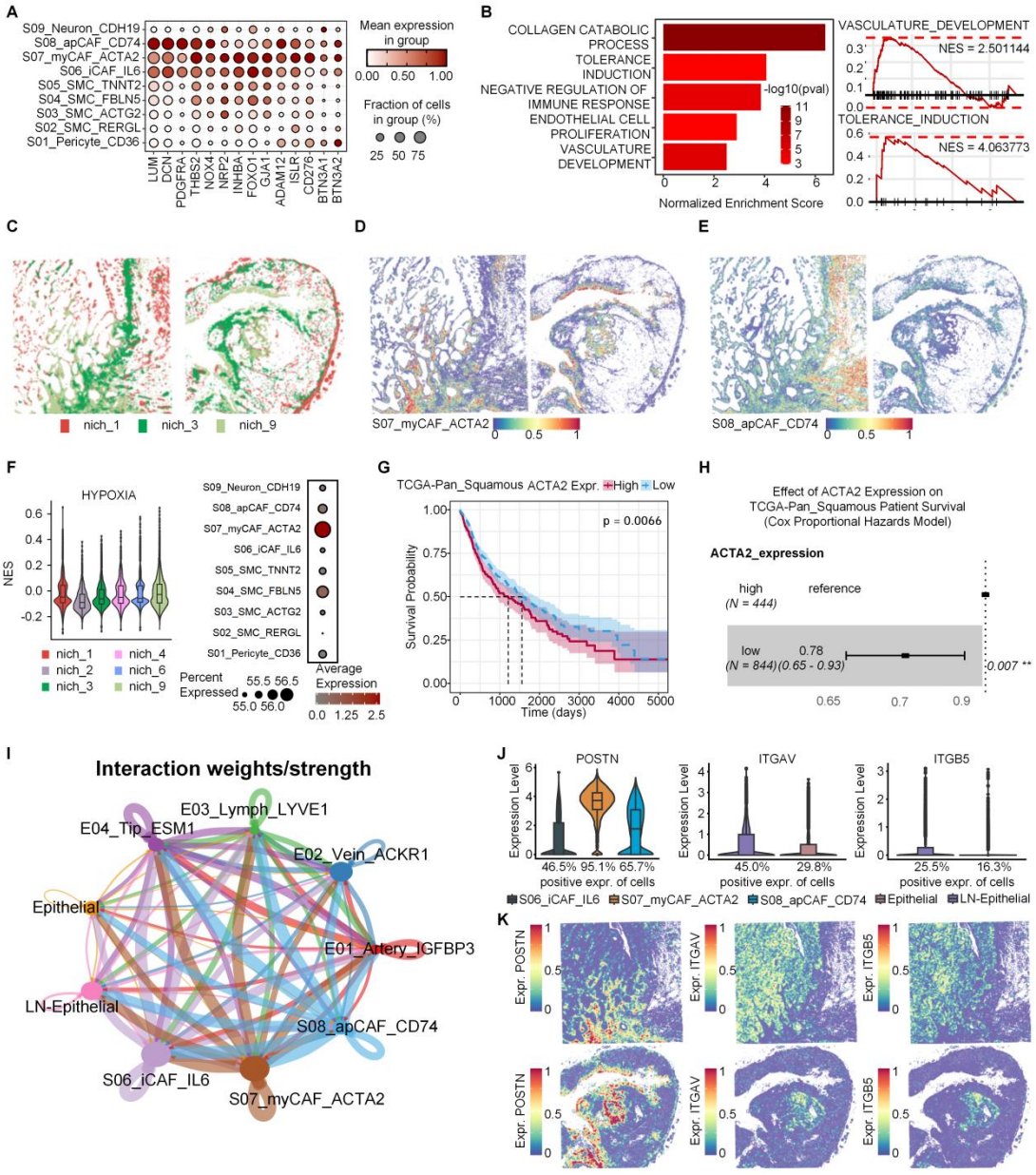

A
S09_Neuron_CDH19
S08_apCAF_CD74
S07_myCAF_ACTA2
S06_iCAF_IL6
S05_SMC_TNNT2
S04_SMC_FBLN5
S03_SMC_ACTG2
S02_SMC_RERGL
S01_Pericyte_CD36
LUM
DCN
PDGFRA
THBS2
NOX4
NRP2
INHBA
FOXO1
GJA1
ADAM12
ISLR
CD276
BTN3A1
BTN3A2
Mean expression in group
0.00 0.50 1.00
Fraction of cells in group (%)
25 50 75
B
COLLAGEN CATABOLIC PROCESS
TOLERANCE INDUCTION
NEGATIVE REGULATION OF IMMUNE RESPONSE
ENDOTHELIAL CELL PROLIFERATION
VASCULATURE DEVELOPMENT
-log10(pval)
Normalized Enrichment Score
VASCULATURE_DEVELOPMENT
NES = 2.501144
TOLERANCE_INDUCTION
NES = 4.063773
C
nich_1
nich_3
nich_9
D
S07_myCAF_ACTA2
E
S08_apCAF_CD74
F
HYPOXIA
NES
nich_1
nich_2
nich_3
nich_4
nich_6
nich_9
Percent Expressed
Average Expression
G
TCGA-Pan_Squamous ACTA2 Expr. High Low
p = 0.0066
Survival Probability
Time (days)
H
Effect of ACTA2 Expression on TCGA-Pan_Squamous Patient Survival (Cox Proportional Hazards Model)
ACTA2_expression
high (N = 444)
reference
low (N = 844)
0.78 (0.65 - 0.93)
0.007 **
I
Interaction weights/strength
E03_Lymph_LYVE1
E04_Tip_ESM1
E02_Vein_ACKR1
Epithelial
E01_Artery_IGFBP3
LN-Epithelial
S08_apCAF_CD74
S06_iCAF_IL6
S07_myCAF_ACTA2
J
POSTN
ITGAV
ITGB5
Expression Level
46.5% 95.1% 65.7%
45.0% 29.8%
25.5% 16.3%
positive expr. of cells
S06_iCAF_IL6
S07_myCAF_ACTA2
S08_apCAF_CD74
Epithelial
LN-Epithelial
K
Expr. POSTN
Expr. ITGAV
Expr. ITGB5

Figure 6

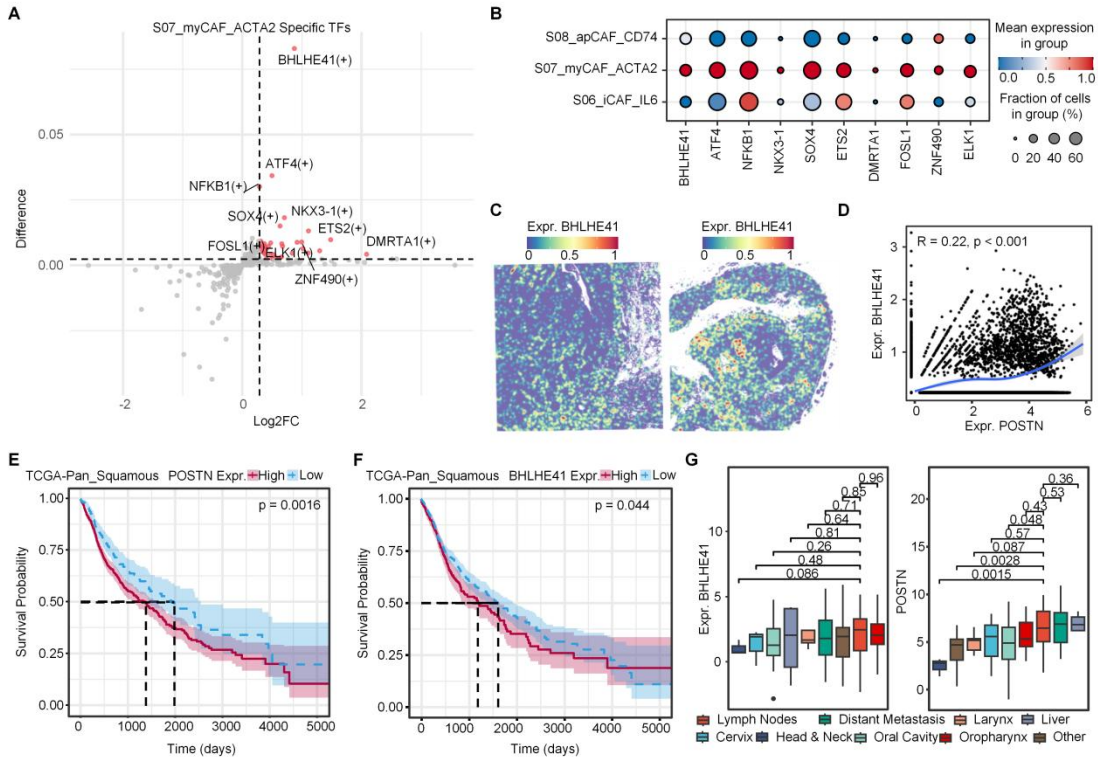
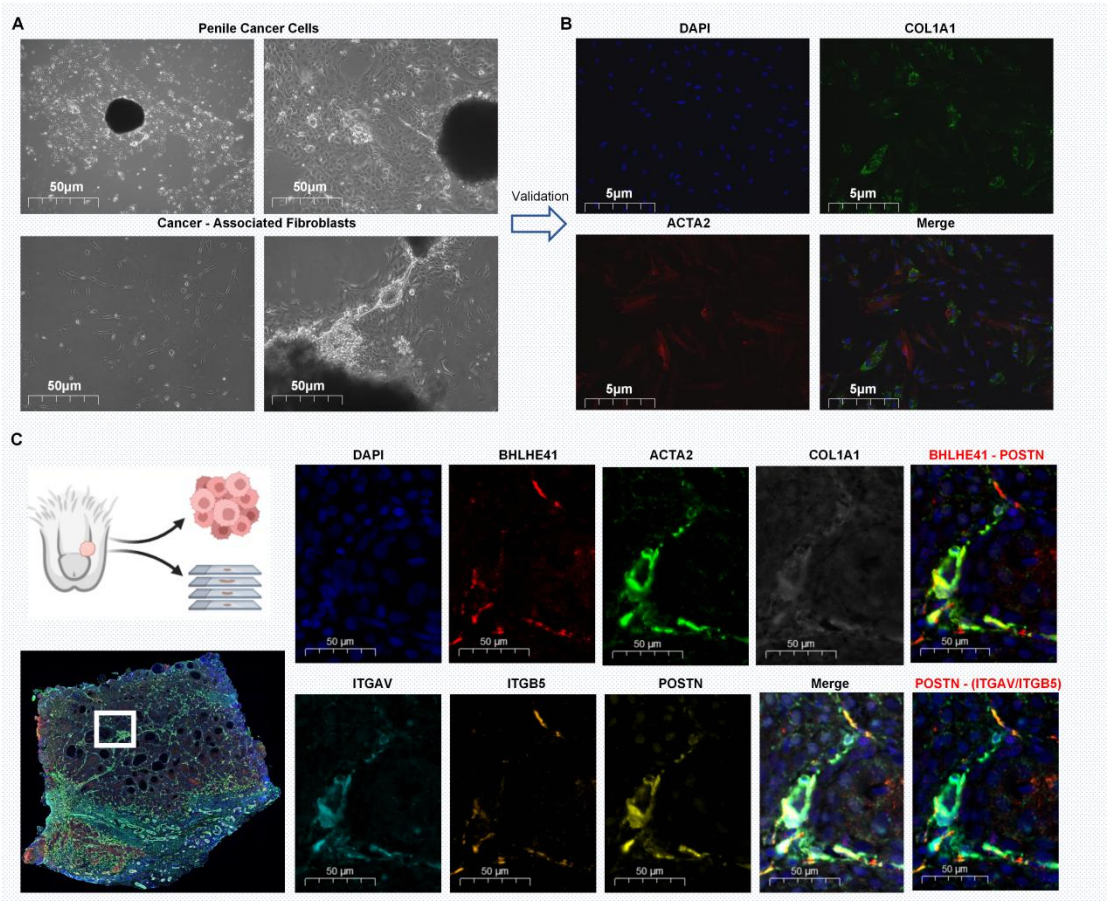
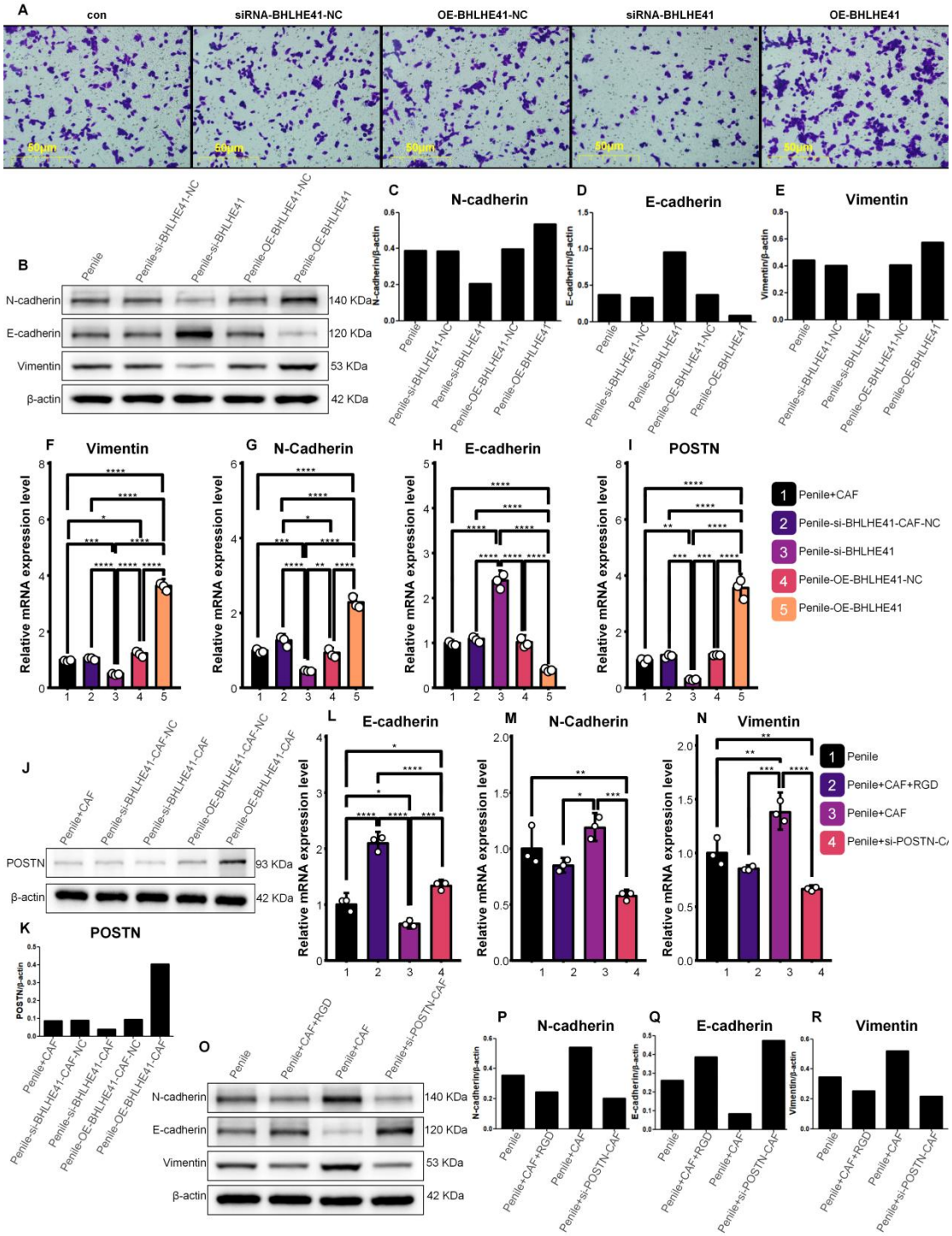

A
S07_myCAF_ACTA2 Specific TFs
BHLHE41(+)
ATF4(+)
NFKB1(+)
SOX4(+)
NKX3-1(+)
ETS2(+)
FOSL1(+)
ELK1(+)
DMRTA1(+)
ZNF490(+)
Difference
Log2FC
B
S08_apCAF_CD74
S07_myCAF_ACTA2
S06_iCAF_IL6
BHLHE41
ATF4
NFKB1
NKX3-1
SOX4
ETS2
DMRTA1
FOSL1
ZNF490
ELK1
Mean expression in group
Fraction of cells in group (%)
C
Expr. BHLHE41
D
R = 0.22, p < 0.001
Expr. BHLHE41
Expr. POSTN
E
TCGA-Pan_Squamous POSTN Expr. High Low
p = 0.0016
Survival Probability
Time (days)
F
TCGA-Pan_Squamous BHLHE41 Expr. High Low
p = 0.044
Survival Probability
Time (days)
G
Expr. BHLHE41
POSTN
Lymph Nodes
Distant Metastasis
Larynx
Liver
Cervix
Head & Neck
Oral Cavity
Oropharynx
Other

Figure 7

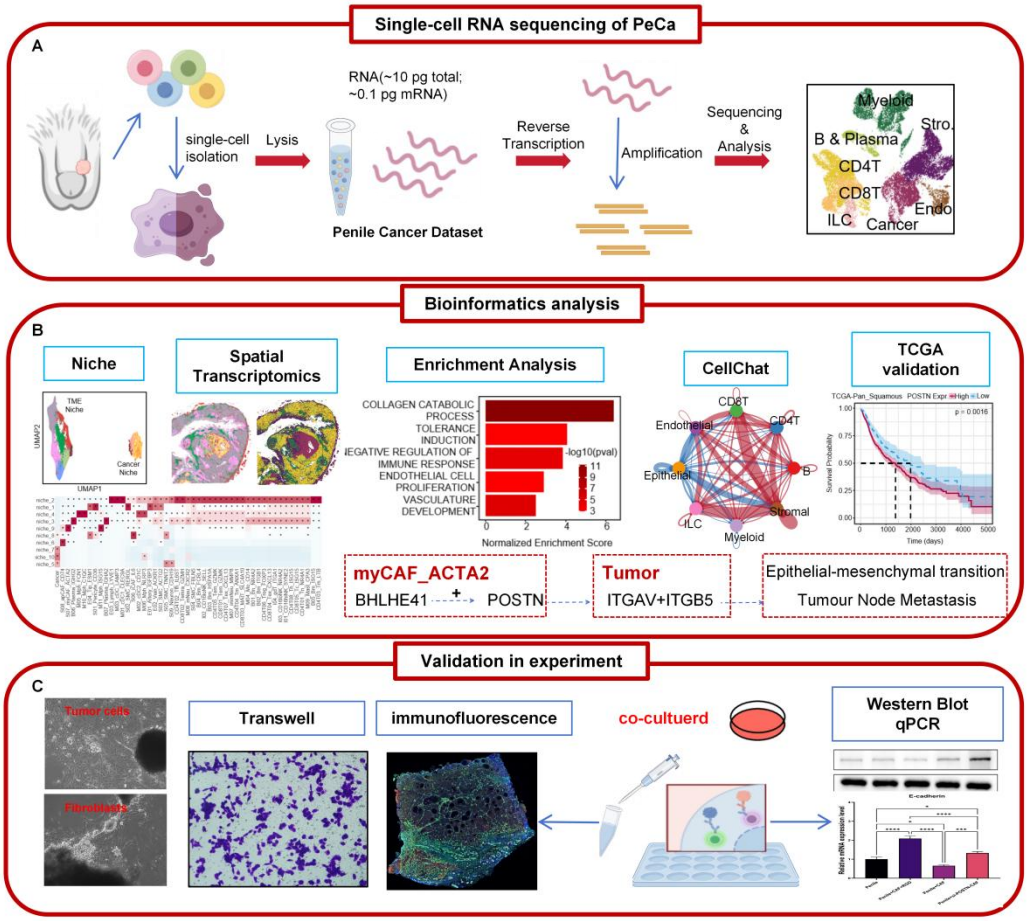

A
Penile Cancer Cells
50µm
50µm
Cancer - Associated Fibroblasts
50µm
50µm
Validation
B
DAPI
COL1A1
5µm
5µm
ACTA2
Merge
5µm
5µm
C
DAPI
BHLHE41
ACTA2
COL1A1
BHLHE41 - POSTN
ITGAV
ITGB5
POSTN
Merge
POSTN - (ITGAV/ITGB5)
50 µm

1007 **Figure 8**



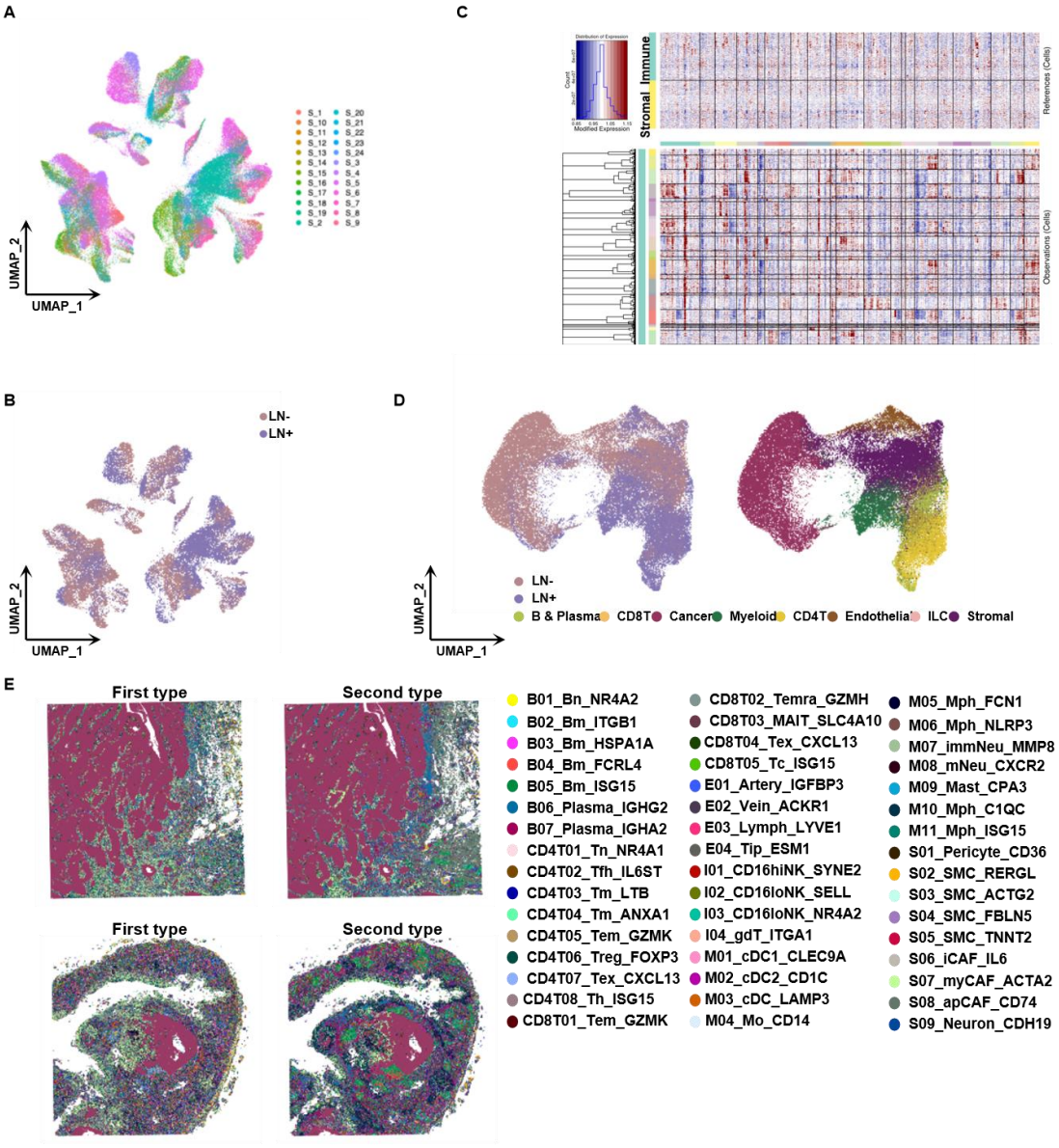
1008

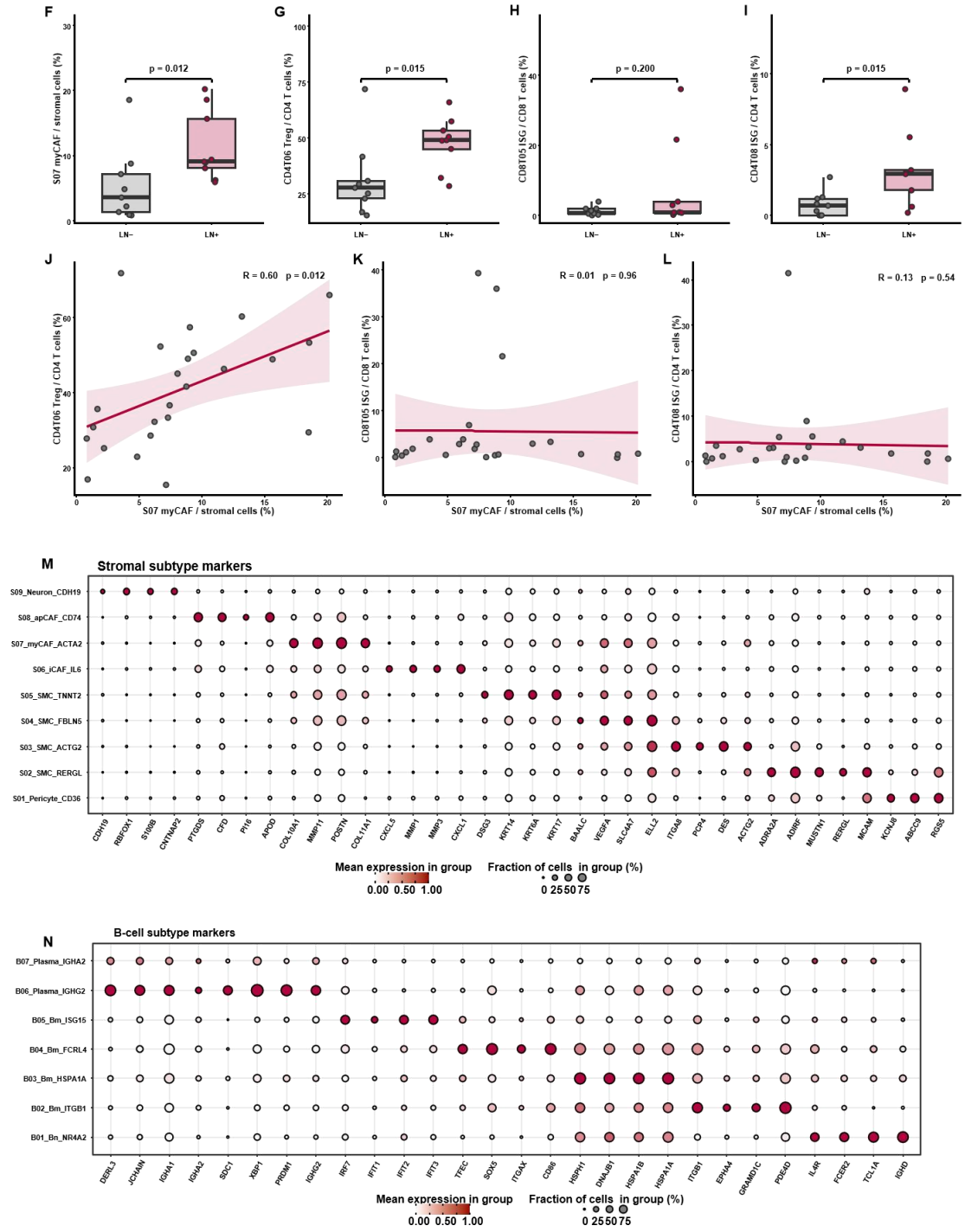
1009 **Scheme 1**



1010

Figure S1


A
UMAP_2
UMAP_1
B
LN-
LN+
C
Immune
Stromal
References (Cells)
Observations (Cells)
D
LN-
LN+
B & Plasma
CD8T
Cancer
Myeloid
CD4T
Endothelia
ILC
Stromal
E
First type
Second type
First type
Second type
B01_Bn_NR4A2
B02_Bm_ITGB1
B03_Bm_HSPA1A
B04_Bm_FCRL4
B05_Bm_ISG15
B06_Plasma_IGHG2
B07_Plasma_IGHA2
CD4T01_Tn_NR4A1
CD4T02_Tfh_IL6ST
CD4T03_Tm_LTB
CD4T04_Tm_ANXA1
CD4T05_Tem_GZMK
CD4T06_Treg_FOXP3
CD4T07_Tex_CXCL13
CD4T08_Th_ISG15
CD8T01_Tem_GZMK
CD8T02_Temra_GZMH
CD8T03_MAIT_SLC4A10
CD8T04_Tex_CXCL13
CD8T05_Tc_ISG15
E01_Artery_IGFBP3
E02_Vein_ACKR1
E03_Lymph_LYVE1
E04_Tip_ESM1
I01_CD16hiNK_SYNE2
I02_CD16loNK_SELL
I03_CD16loNK_NR4A2
I04_gdT_ITGA1
M01_cDC1_CLEC9A
M02_cDC2_CD1C
M03_cDC_LAMP3
M04_Mo_CD14
M05_Mph_FCN1
M06_Mph_NLRP3
M07_immNeu_MMP8
M08_mNeu_CXCR2
M09_Mast_CPA3
M10_Mph_C1QC
M11_Mph_ISG15
S01_Pericyte_CD36
S02_SMC_RERGL
S03_SMC_ACTG2
S04_SMC_FBLN5
S05_SMC_TNNT2
S06_iCAF_IL6
S07_myCAF_ACTA2
S08_apCAF_CD74
S09_Neuron_CDH19


F
S07 myCAF / stromal cells (%)
p = 0.012
LN−
LN+
G
CD4T06 Treg / CD4 T cells (%)
p = 0.015
H
CD8T05 ISG / CD8 T cells (%)
p = 0.200
I
CD4T08 ISG / CD4 T cells (%)
p = 0.015
J
R = 0.60 p = 0.012
S07 myCAF / stromal cells (%)
K
CD8T05 ISG / CD8 T cells (%)
R = 0.01 p = 0.96
L
CD4T08 ISG / CD4 T cells (%)
R = 0.13 p = 0.54
M
Stromal subtype markers
S09_Neuron_CDH19
S08_apCAF_CD74
S07_myCAF_ACTA2
S06_iCAF_IL6
S05_SMC_TNNT2
S04_SMC_FBLN5
S03_SMC_ACTG2
S02_SMC_RERGL
S01_Pericyte_CD36
CDH19 RBFOX1 S100B CNTNAP2 PTGDS CFD PI16 APOD COL10A1 MMP11 POSTN COL11A1 CXCL5 MMP1 MMP3 CXCL1 DSG3 KRT14 KRT6A KRT17 BAALC VEGFA SLC4A7 ELL2 ITGA6 PCP4 DES ACTG2 ADRA2A ADIRF MUSTN1 RERGL MCAM KCNJ8 ABCC9 RGS5
Mean expression in group
0.00 0.50 1.00
Fraction of cells in group (%)
0 25 50 75
N
B-cell subtype markers
B07_Plasma_IGHA2
B06_Plasma_IGHG2
B05_Bm_ISG15
B04_Bm_FCRL4
B03_Bm_HSPA1A
B02_Bm_ITGB1
B01_Bn_NR4A2
DERL3 JCHAIN IGHA1 IGHA2 SDC1 XBP1 PRDM1 IGHG2 IRF7 IFIT1 IFIT2 IFIT3 TFEC SOX5 ITGAX CD96 HSPH1 DNAJB1 HSPA1B HSPA1A ITGB1 EPHA4 GRAMD1C PDE4D IL4R FCER2 TCL1A IGHD
Mean expression in group
0.00 0.50 1.00
Fraction of cells in group (%)
0 25 50 75

1013

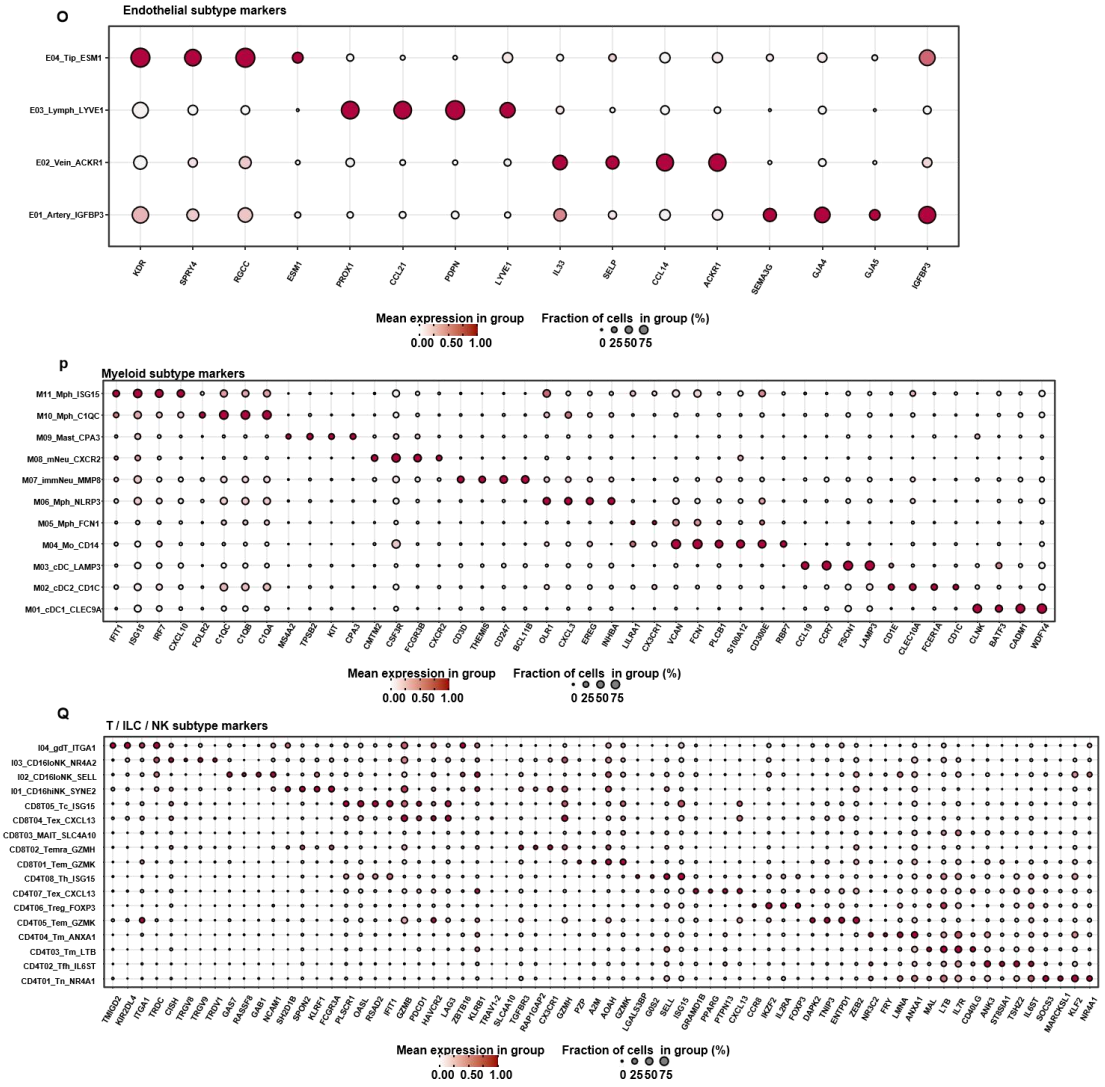
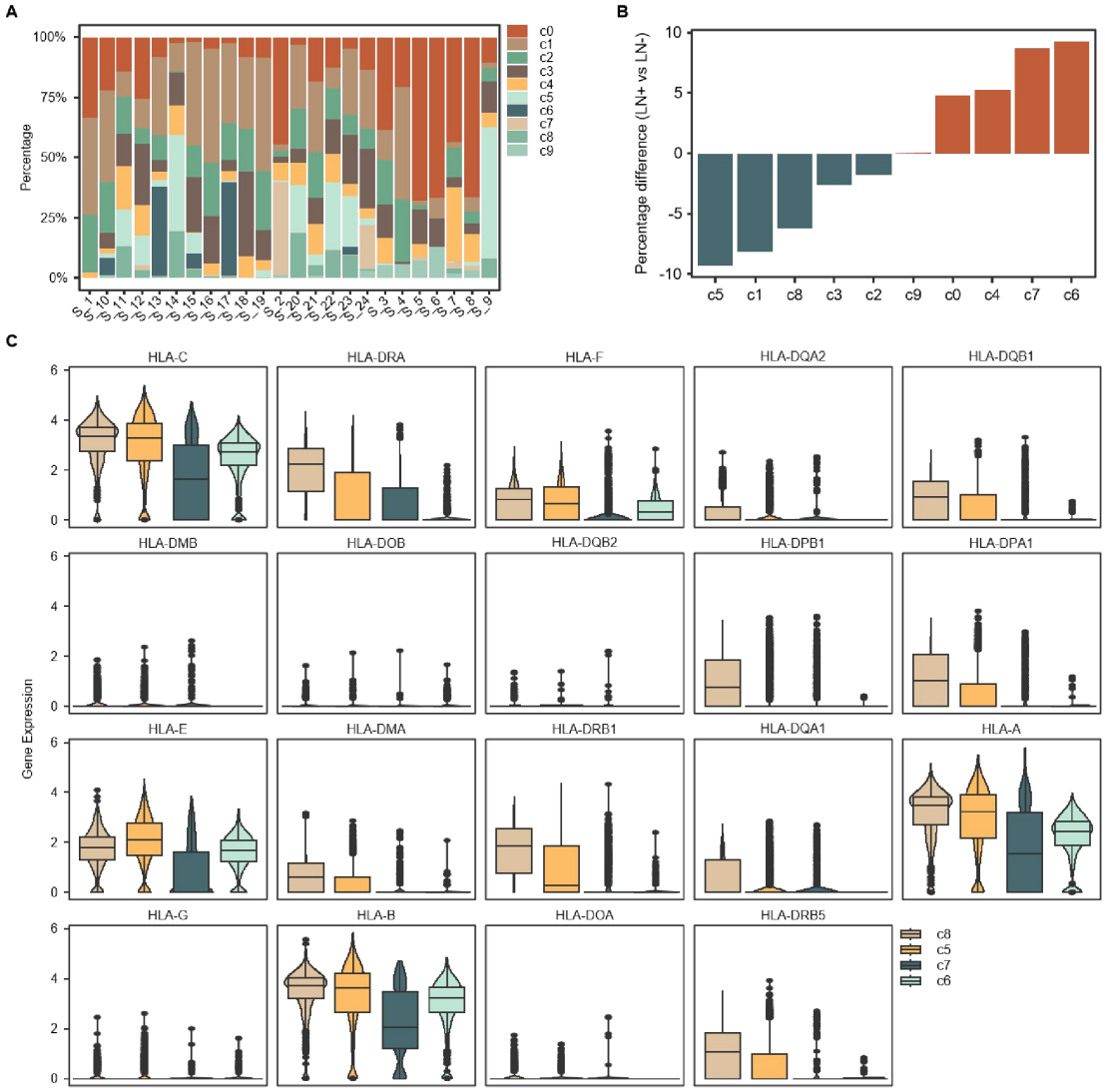
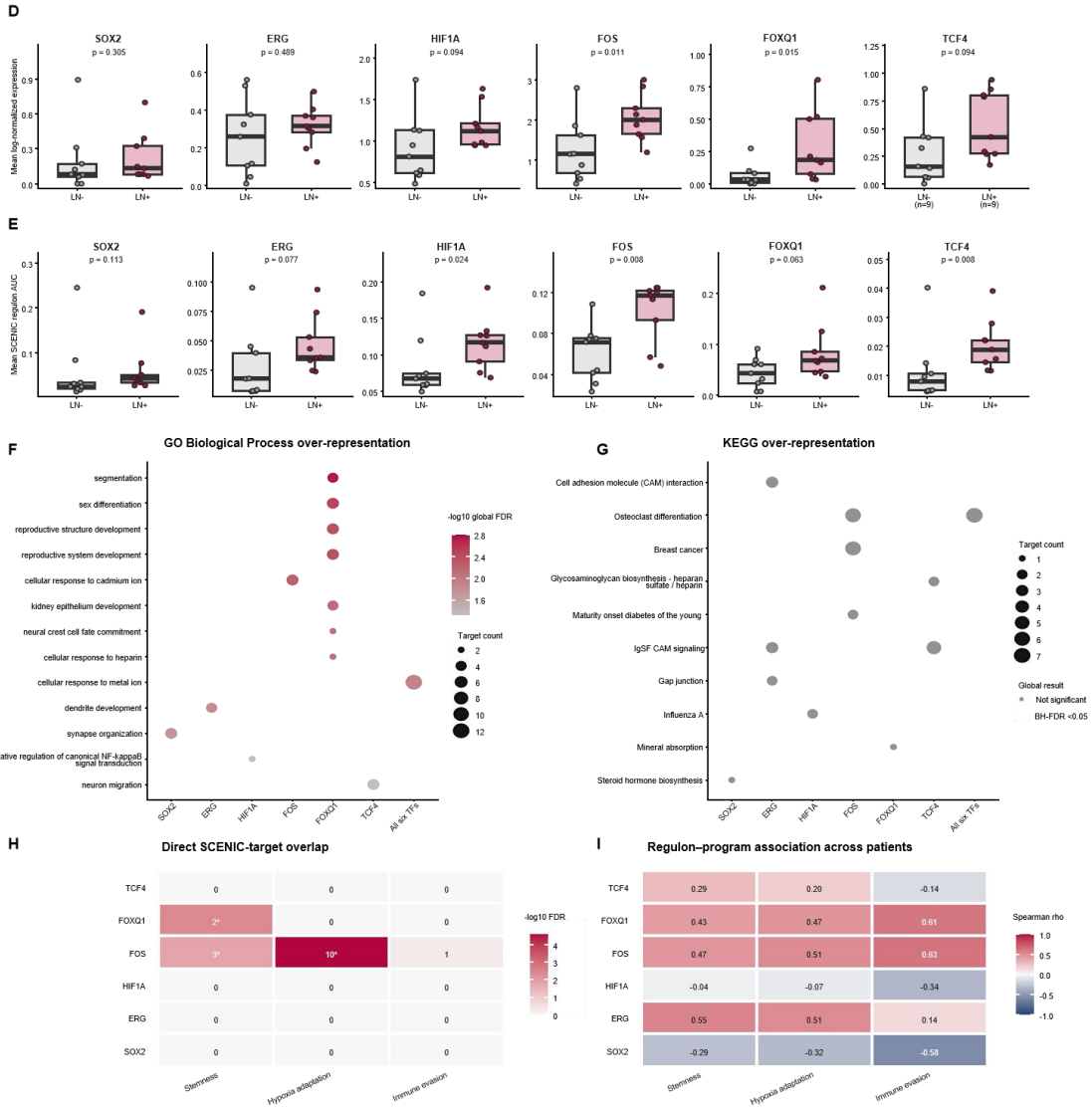

o
Endothelial subtype markers
E04_Tip_ESM1
E03_Lymph_LYVE1
E02_Vein_ACKR1
E01_Artery_IGFBP3
KDR
SPRY4
RGCC
ESM1
PROX1
CCL21
PDPN
LYVE1
IL33
SELP
CCL14
ACKR1
SEMA3G
GJA4
GJA5
IGFBP3
Mean expression in group
0.00 0.50 1.00
Fraction of cells in group (%)
0 25 50 75
p
Myeloid subtype markers
M11_Mph_ISG15
M10_Mph_C1QC
M09_Mast_CPA3
M08_mNeu_CXCR2
M07_immNeu_MMP8
M06_Mph_NLRP3
M05_Mph_FCN1
M04_Mo_CD14
M03_cDC_LAMP3
M02_cDC2_CD1C
M01_cDC1_CLEC9A
IFIT1
ISG15
IRF7
CXCL10
FOLR2
C1QC
C1QB
C1QA
MS4A2
TPSB2
KIT
CPA3
CMTM2
CSF3R
FCGR3B
CXCR2
CD33
THEMIS
CD247
BCL11B
OLR1
CXCL3
ERG
MMRA
LILRA1
CXCR1
VCAN
FCN1
PLCB1
S100A12
CD300E
RBP7
CCL19
CCR7
FSCN1
LAMP3
CD1E
CLEC10A
FCER1A
CD1C
CLNK
BATF3
CADM1
WDFY4
Mean expression in group
0.00 0.50 1.00
Fraction of cells in group (%)
0 25 50 75
Q
T / ILC / NK subtype markers
I04_gdT_ITGA1
I03_CD16loNK_NR4A2
I02_CD16loNK_SELL
I01_CD16hiNK_SYNE2
CD8T05_Tc_ISG15
CD8T04_Tex_CXCL13
CD8T03_MAIT_SLC4A10
CD8T02_Temra_GZMH
CD8T01_Tem_GZMK
CD4T08_Th_ISG15
CD4T07_Tex_CXCL13
CD4T06_Treg_FOXP3
CD4T05_Tem_GZMK
CD4T04_Tm_ANXA1
CD4T03_Tm_LTB
CD4T02_Tfh_IL6ST
CD4T01_Tn_NR4A1
TMIGD2
KIR2DL4
ITGA1
TRDC
CISH
TTC9B
TTC9
TYROV1
GAS7
RASSF8
GAB1
NCAM1
SH2D1B
SPON2
KLRF1
FCGR3A
PLSCR1
OASL
RSAD2
IFIT1
GZMB
PDCD1
HAVCR2
LAG3
ZBTB16
KLRB1
TIGIT-2
SLC4A10
TOP3B3
RAPTGRP2
CXCR1
CD4H
P2P
ADM
AOAH
GZMK
LGALS3BP
GIB2
SELL
ISG15
GRAMD1B
PPARG
PTPN13
CXCL13
CCR8
IKZF2
IL2RA
FOXP3
DAPK2
TNP3
ENTPD1
ZEB2
NR3C2
FRY
LMNA
ANXA1
MAL
LTB
IL7R
CD40LG
ANK3
STBSA1
TSHZ2
IL6ST
SOCS3
MARCHCS1
KLF2
NR4A1
Mean expression in group
0.00 0.50 1.00
Fraction of cells in group (%)
0 25 50 75
R
MHC class I and II expression across epithelial subtypes
c5
c8
c1
c3
c2
c9
c0
c4
c6
c7
HLA-DRA
HLA-DQA2
HLA-DQB1
HLA-DMB
HLA-DOB
HLA-DQB2
HLA-DPB1
HLA-DPA1
HLA-DMA
HLA-DRB1
HLA-DQA1
HLA-DOA
HLA-DRB5
Mean expression in group
0.00 0.50 1.00
Fraction of cells in group (%)
0 25 50 75

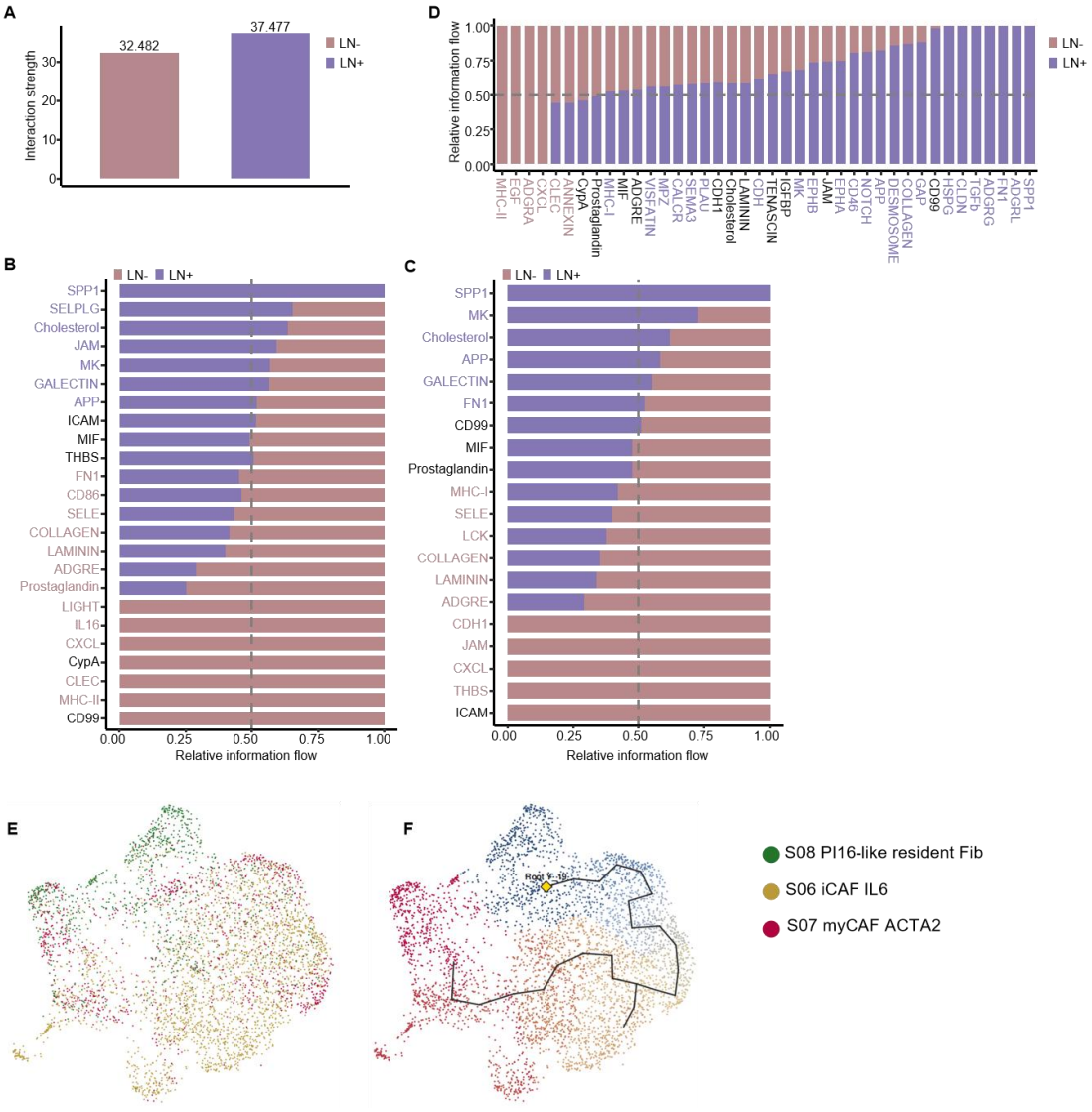
Figure S2

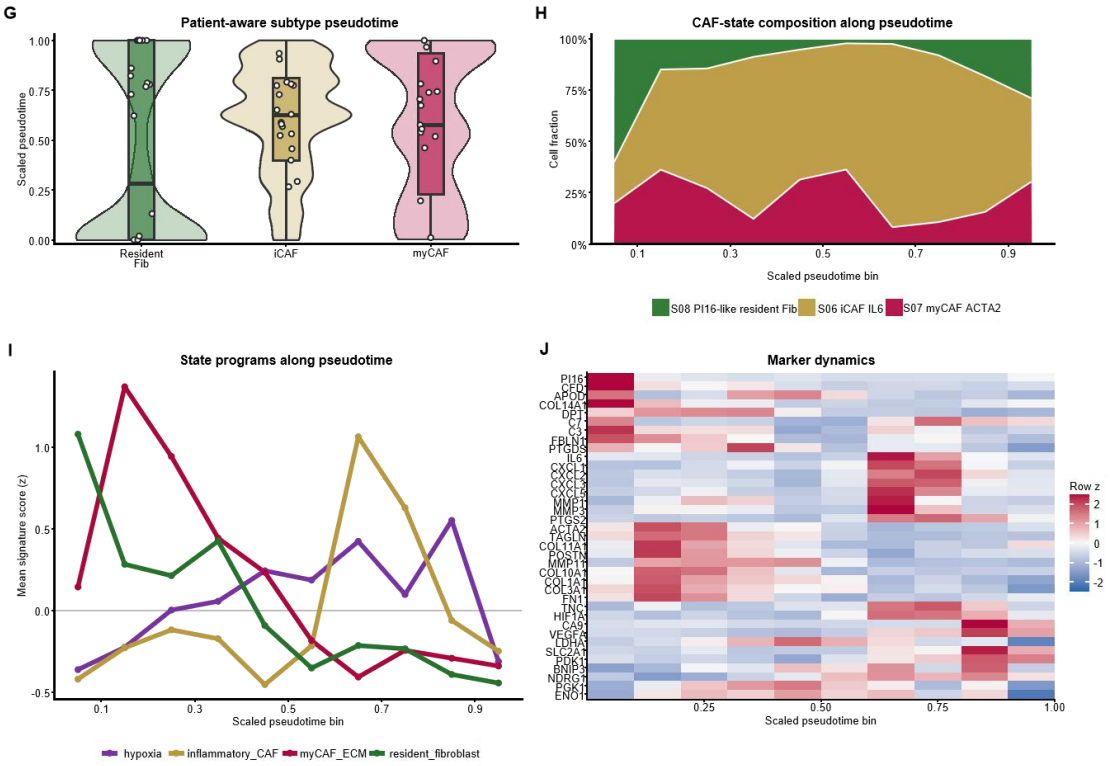

A
100%
75%
50%
25%
0%
Percentage
c0
c1
c2
c3
c4
c5
c6
c7
c8
c9
B
Percentage difference (LN+ vs LN-)
10
5
0
-5
-10
c5
c1
c8
c3
c2
c9
c0
c4
c7
c6
C
HLA-C
HLA-DRA
HLA-F
HLA-DQA2
HLA-DQB1
HLA-DMB
HLA-DOB
HLA-DQB2
HLA-DPB1
HLA-DPA1
HLA-E
HLA-DMA
HLA-DRB1
HLA-DQA1
HLA-A
HLA-G
HLA-B
HLA-DOA
HLA-DRB5
Gene Expression
c8
c5
c7
c6


D
SOX2
p = 0.305
ERG
p = 0.489
HIF1A
p = 0.094
FOS
p = 0.011
FOXQ1
p = 0.015
TCF4
p = 0.094
Mean log-normalised expression
LN-
LN+
E
SOX2
p = 0.113
ERG
p = 0.077
HIF1A
p = 0.024
FOS
p = 0.008
FOXQ1
p = 0.063
TCF4
p = 0.008
Mean SCENIC regulon AUC
F
GO Biological Process over-representation
segmentation
sex differentiation
reproductive structure development
reproductive system development
cellular response to cadmium ion
kidney epithelium development
neural crest cell fate commitment
cellular response to heparin
cellular response to metal ion
dendrite development
synapse organization
negative regulation of canonical NF-kappaB signal transduction
neuron migration
-log10 global FDR
Target count
G
KEGG over-representation
Cell adhesion molecule (CAM) interaction
Osteoclast differentiation
Breast cancer
Glycosaminoglycan biosynthesis - heparan sulfate / heparin
Maturity onset diabetes of the young
IgSF CAM signaling
Gap junction
Influenza A
Mineral absorption
Steroid hormone biosynthesis
Target count
Global result
Not significant
BH-FDR <0.05
SOX2
ERG
HIF1A
FOS
FOXQ1
TCF4
All six TFs
H
Direct SCENIC-target overlap
TCF4
FOXQ1
FOS
HIF1A
ERG
SOX2
-log10 FDR
Stemness
Hypoxia adaptation
Immune evasion
I
Regulon–program association across patients
Spearman rho

1017

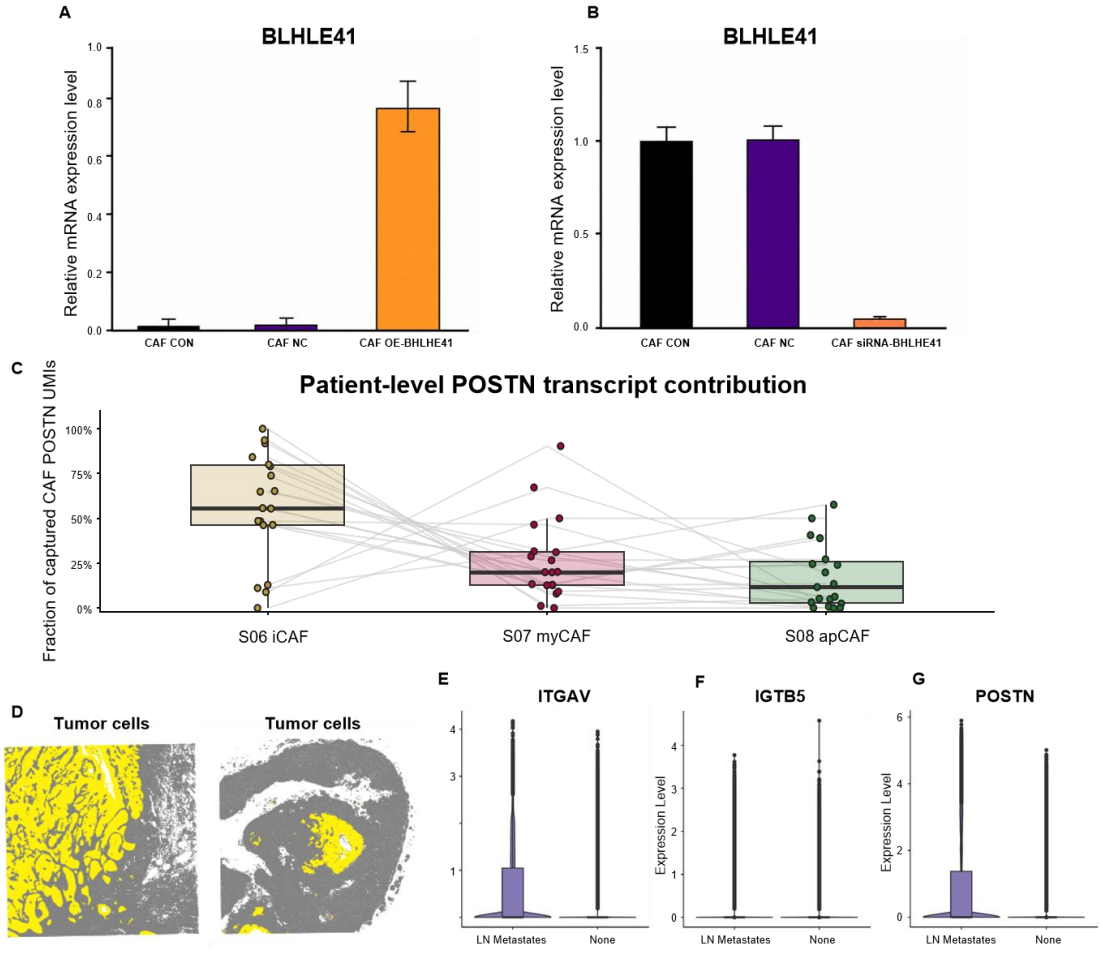
Figure S3

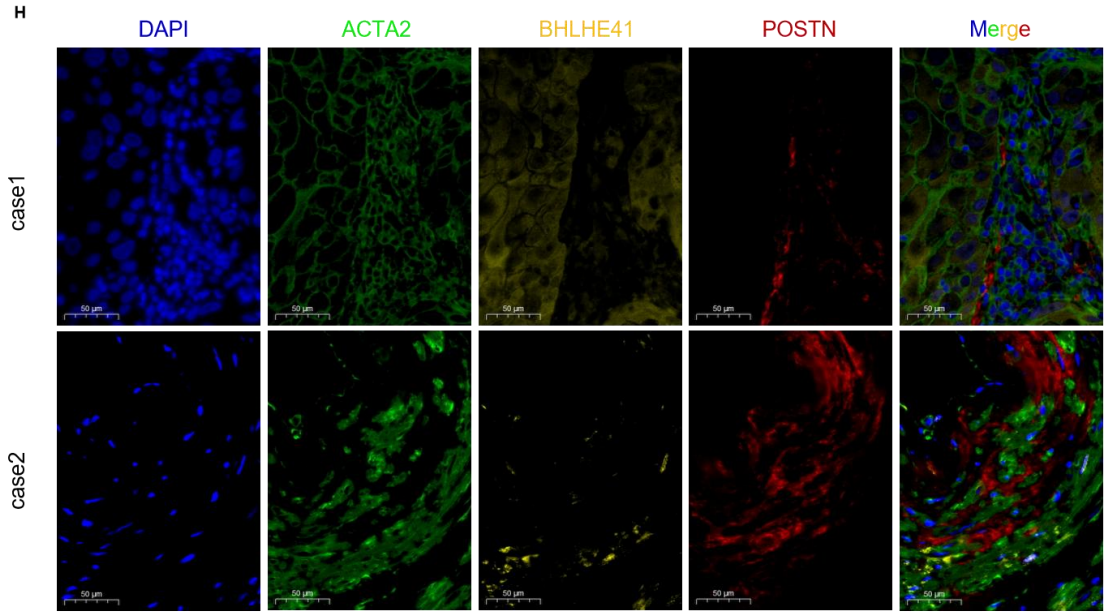

A
32.482
37.477
LN-
LN+
Interaction strength
B
C
Relative information flow
D
Relative information flow
E
F
S08 PI16-like resident Fib
S06 iCAF IL6
S07 myCAF ACTA2

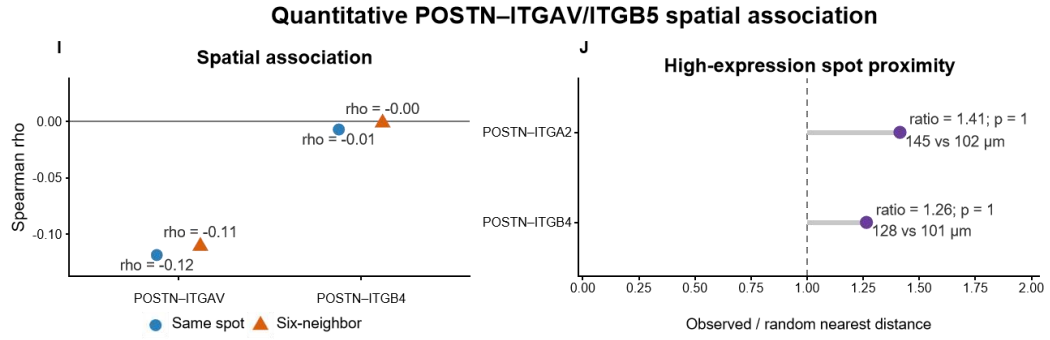

G
Patient-aware subtype pseudotime
Scaled pseudotime
Resident Fib
iCAF
myCAF
H
CAF-state composition along pseudotime
Cell fraction
Scaled pseudotime bin
S08 PI16-like resident Fib
S06 iCAF IL6
S07 myCAF ACTA2
I
State programs along pseudotime
Mean signature score (z)
Scaled pseudotime bin
hypoxia
inflammatory_CAF
myCAF_ECM
resident_fibroblast
J
Marker dynamics
Row z
Scaled pseudotime bin

1020

Figure S4


A
BLHLE41
Relative mRNA expression level
CAF CON
CAF NC
CAF OE-BHLHE41
B
BLHLE41
Relative mRNA expression level
CAF CON
CAF NC
CAF siRNA-BHLHE41
C
Patient-level POSTN transcript contribution
Fraction of captured CAF POSTN UMIs
S06 iCAF
S07 myCAF
S08 apCAF
D
Tumor cells
Tumor cells
E
ITGAV
F
IGTB5
G
POSTN
Expression Level
LN Metastates
None


H
DAPI
ACTA2
BHLHE41
POSTN
Merge
case1
case2


Quantitative POSTN–ITGAV/ITGB5 spatial association
I
Spatial association
Spearman rho
rho = -0.00
rho = -0.01
rho = -0.11
rho = -0.12
POSTN–ITGAV
POSTN–ITGB4
Same spot
Six-neighbor
J
High-expression spot proximity
POSTN–ITGA2
POSTN–ITGB4
ratio = 1.41; p = 1
145 vs 102 µm
ratio = 1.26; p = 1
128 vs 101 µm
Observed / random nearest distance

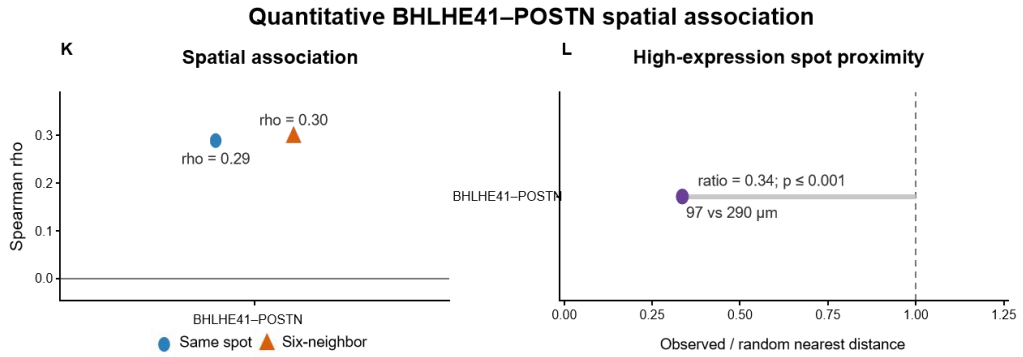

Quantitative BHLHE41–POSTN spatial association
K
Spatial association
Spearman rho
rho = 0.30
rho = 0.29
BHLHE41–POSTN
Same spot
Six-neighbor
L
High-expression spot proximity
BHLHE41–POSTN
ratio = 0.34; p ≤ 0.001
97 vs 290 µm
Observed / random nearest distance

Figure Legends

Figure 1 | Single-cell transcriptomic profiling of 24 penile cancer patients and spatial transcriptomic analysis of one representative case.

A, Uniform Manifold Approximation and Projection (UMAP) plot showing the major cell types and their subtypes in samples across the LN+ and LN- groups, with individual cells represented as dots and different cell populations distinguished by colors. Cell annotation: Tn, naïve T cells; Tfh, follicular helper T cell; Tm, memory T cell; Tem effector memory cells; Treg, regulatory T cells; Tex, exhausted T cells; Th, T helper cells; Temra, terminally differentiated effector memory or effector; Tc, Cytotoxic T cells; MAIT, Mucosal-Associated Invariant T cell; Bm, B memory cells; Bn, naïve B cells; NK, natural killer cells; Artery cells; Vein cells; Lymph cells; Tip cells; gdT, Gamma Delta T Cell; cDC, Conventional Dendritic Cells; Mo, Monocytes; Mph, Macrophages; immNeu, immature neutrophils; mNeu, mature neutrophils; Mast cells; Pericytes; SMC, smooth muscle cells; iCAF, inflammatory cancer-associated fibroblast (CAF); myCAF, myofibroblastic CAF; apCAF, Antigen-presenting CAF; Neuron cells.

B, Dot plot shows the expression of signature genes of each PeCa cell type. Color indicates the expression level and size indicates the expression fraction.

C, Spatial mapping of Visium HD, colored by cell types as in A.

D, Box plots showing the patient-level proportions of major cell types in the LN+ and LN− groups. Each point represents one patient. Statistical comparisons were performed using a two-sided Student's t-test.

E, The bar chart illustrates the changes in proportions of different cell types relative to total cells in the metastatic group. Color represents subtypes of CD4 T cells, and length represents the magnitude of change.

F, The bar chart illustrates the changes in proportions of different cell types relative to total cells in the metastatic group. Color represents subtypes of B cells, and length represents the magnitude of change.

G, The bar chart illustrates the changes in proportions of different cell types relative to total cells in the metastatic group. Color represents subtypes of stromal cells, and length represents the magnitude of change.

Figure 2 | Subtypes of epithelial cells in PeCa.

A, UMAP plot showing distribution of tumor cells by the LN+ and LN- groups (left) and subtypes of tumor cells (right).

B, Bar plot showing differences in the distribution of tumor epithelial subtypes between the LN+ and LN- groups, with bar length representing the Odds Ratio (OR) values of each subtype.

C, Violin plot showing stemness scores in different subtypes of tumor cells.

D, Bar plot shows enrichment results of downregulated genes in the LN+ group.

E, Enrichment of gene signature based on GSEA analysis between the LN+ and LN- groups.

F, Dot plot displays expression levels of specific MHC class I and class II genes in C5-C8 subtypes, where dot size represents the fraction of cells in each group, and dot color indicates the mean expression level within the group.

G, Relative expression levels of genes of interest in different groups.

H, The dot plot displays the activity of transcription factors in the LN+ group, with red representing transcription factors specifically activated in the LN+ group.

I, UMAP plots show the expression of specifically activated transcription factors SOX2, ERG,

HIF1A, FOS, FOXQ1 and TCF4.

J, Violin plot shows hypoxia scores between the LN+ and LN- groups.

K, Spatial mapping of Visium HD, with spots colored by hypoxia score.

L, The bar plot displays the normalized score of 15 significantly upregulated genes in the early lymph node metastasis group.

Figure 3 | Cell-cell communication plots in the microenvironment of PeCa

A-B, Dot plots show the cell-cell communication signal strength between different cell types, with the ordinate representing the incoming signal strength and the abscissa representing the outgoing signal strength. Dots represent cell types, distinguished by different colors.

C, Circle diagrams show cell-cell communication among cell types in the LN+ and LN- groups. Red represents upregulated signal strength.

D, Dot plots show the differential cell-cell communication signal strength of different signal pathways in different cell types, with the ordinate representing the differential incoming signal strength and the abscissa representing the differential outgoing signal strength. Dots represent signaling pathways, distinguished by different colors.

E-F, Bar plots show differences in the distribution of subtypes of T cells and stromal cells or endothelial cells between the LN+ and LN- groups, with bar length representing the Odds Ratio (OR) values of each subtype.

G, Dot plot displays expression levels of immunosuppressive molecules in stromal cell subtypes, where dot size represents the fraction of cells in each group, and dot color indicates the mean expression level within the group.

Figure 4 | Cellular niche composition in PeCa.

A, Heatmap showing scaled median cell-type compositions within each niche. Asterisks indicate increased abundance of a cell type in a niche compared with other niches (one-sided Wilcoxon rank sum test, adjusted (adj.) $P < 0.05$).

B, UMAP plot showing distribution of all niches.

C, The spatial heatmap highlights the distribution of all niches in primary lesions and matched lymph node metastases.

D, The bar chart shows the proportion of all niches in primary lesions and matched lymph node metastases.

E, Bar plot showing differences in the distribution of all niches in primary lesions and matched lymph node metastases, with bar length representing the Odds Ratio (OR) values of each subtype.

Figure 5 | The relationship of hypoxic myCAF subtypes and the progression of PeCa.

A, Dot plot displays expression levels of marker genes of S07_myCAF_ACTA2 in stromal subtypes, where dot size represents the fraction of cells in each group, and dot color indicates the mean expression level within the group.

B, Enrichment of marker genes of S07_myCAF_ACTA2 based on GSEA analysis.

C, The spatial heatmap highlights the distribution of niche_1,3 and 9 in primary lesions and matched lymph node metastases.

D-E, The spatial heatmap highlights the distribution of myCAF and apCAF in primary lesions and matched lymph node metastases.

F, Violin plot showing hypoxia scores across six of the ten cellular niches defined in Figure 4: niche_1, niche_2, niche_3, niche_4, niche_6, and niche_9 (left). Dot plot displays hypoxia scores in stromal subtypes, where dot size represents the fraction of cells in each group, and dot color

indicates the mean expression level within the group (right).

G, Survival curve shows the relationship between expression levels of ACTA2 and prognosis in TCGA-Pan_Squamous. Red indicates high expression, whereas blue indicates low expression.

H, Forest plot showing the association between ACTA2 expression and survival in TCGA-Pan_Squamous using a Cox proportional hazards model.

I, CellChat analysis showing the interaction strength between CAF subtypes and endothelial cell subtypes in the non-metastatic and metastatic groups. Edge width represents the inferred interaction strength.

J, Violin plots showing POSTN expression across CAF subtypes and ITGAV/ITGB5 expression across epithelial cell populations.

K, Spatial expression maps showing the distribution of POSTN, ITGAV and ITGB5 in the primary tumor and matched lymph-node metastasis from the representative spatially profiled case.

Figure 6 | Mechanistic analysis of myCAF in promoting lymph node metastasis of PeCa.

A, The scatter plot displays the activity of transcription factors in S07_myCAF_ACTA2, with red representing transcription factors specifically activated in S07_myCAF_ACTA2.

B, Dot plot displays expression levels of S07_myCAF_ACTA2 specific TFs, where dot size represents the fraction of cells in each group, and dot color indicates the mean expression level within the group.

C, The spatial heatmap highlights the distribution of BHLHE41 expression in primary lesions and matched lymph node metastases.

D, Scatter plot depicts the expression levels of POSTN and BHLHE41, along with their correlation.

E-F, Survival curves show the relationship between expression levels of POSTN or BHLHE41 and prognosis in TCGA-Pan_Squamous. Red indicates high expression, whereas blue indicates low expression.

G, Expression distributions of BHLHE41 and POSTN across different tissue/sample-site categories.

Figure 7 | Culture of penile cancer cells and immunofluorescence staining.

A, Growth patterns of penile cancer cells and penile cancer-associated fibroblasts.

B, Representative immunofluorescence images of cultured CAFs stained for COL1A1 and ACTA2, with DAPI nuclear counterstaining.

C, Multiplex immunofluorescence imaging of the penile cancer tumor microenvironment. DAPI (blue), BHLHE41 (red), ACTA2 (green), COL1A1 (white), ITGAV (cyan), ITGB5 (orange), and POSTN (yellow).

Figure 8 | Validation of the mechanism of the BHLHE41-POSTN-(ITGAV+ITGB5) axis in PeCa.

A, Transwell assay results demonstrating the invasive capacity of penile cancer cells under different experimental groups.

B-E, Western blot analysis showing changes in the protein expression of EMT markers N-cadherin, E-cadherin, and Vimentin under conditions of BHLHE41 overexpression or knockdown. Co-cultured cells were collected, lysed, and centrifuged, and the resulting cell lysate supernatants were used for Western blot analysis. Panel B shows the representative Western blot images, while panels

C-E show the corresponding quantitative analyses of N-cadherin (C), E-cadherin (D), and Vimentin (E), respectively.

F-H, qPCR analysis showing the transcriptional levels of Vimentin (F), N-cadherin (G) and E-cadherin (H), in cells collected after co-culture under conditions of BHLHE41 overexpression or knockdown.

I, qPCR analysis showing the transcriptional level of POSTN in cells collected after co-culture under conditions of BHLHE41 overexpression or knockdown.

J-K, Western blot analysis showing POSTN protein expression under different co-culture conditions with BHLHE41 overexpression or knockdown. Co-cultured cells were lysed and centrifuged, and the resulting cell lysate supernatants were used for protein detection. Panel J shows the representative Western blot images, while panel K shows the corresponding quantitative analysis of POSTN protein expression.

L-N, qPCR analysis showing the transcriptional levels of EMT markers in PeCa cells under four experimental conditions: PeCa cells alone, PeCa cells co-cultured with CAFs in the presence of the integrin-blocking RGD peptide, PeCa cells co-cultured with untreated CAFs, and PeCa cells co-cultured with POSTN-knockdown CAFs. Following co-culture, PeCa cells were separately harvested for RNA extraction and qPCR analysis. Panels L-N show the transcriptional levels of E-cadherin (L), N-cadherin (M), and Vimentin (N), respectively.

O-R, Western blot analysis showing changes in the protein expression levels of the EMT markers N-cadherin, E-cadherin, and Vimentin in PeCa cells under four experimental conditions: PeCa cells alone, PeCa cells co-cultured with CAFs in the presence of the integrin-blocking RGD peptide, PeCa cells co-cultured with untreated CAFs, and PeCa cells co-cultured with POSTN-knockdown CAFs. Following co-culture, PeCa cells were separately harvested for protein analysis. Panel O shows the representative Western blot images, while panels P-R show the corresponding quantitative analyses of N-cadherin (P), E-cadherin (Q), and Vimentin (R), respectively.

Scheme 1 | Schematic overview of the study design and proposed mechanism underlying lymph node metastasis in penile cancer.

Single-cell RNA sequencing and spatial transcriptomic analyses were integrated to characterize the cellular landscape of penile cancer (PeCa) and identify metastasis-associated cellular states and spatial niches. Bioinformatic analyses identified a hypoxia-adapted myCAF_ACTA2 population enriched in metastasis-associated niches and highlighted BHLHE41-associated POSTN signaling as a candidate stromal mechanism linked to tumor progression. Cell-cell communication analysis suggested a potential interaction between myCAF_ACTA2-derived POSTN and ITGAV/ITGB5 on tumor cells. Primary-cell co-culture, Transwell assays, qPCR, Western blotting, and multiplex immunofluorescence provided in vitro support for an EMT-associated, pro-invasive phenotype, while the in situ spatial organization of this signaling relationship remains to be established.

Figure S1 | Single-cell RNA-seq and spatial transcriptomic analyses.

A, UMAP plot of all scRNA-seq cells colored by patient/sample identifier.

B, UMAP plot shows the distribution of PeCa cells by the LN+ and LN- groups in all samples.

C, InferCNV was performed separately on PeCa, and cancer cells were grouped according to epithelial cells, with stromal cells and immune cells as reference populations.

D, UMAP plot shows the distribution of PeCa cells by cell types (left) or LN+ and LN- groups (right) in primary lesions and matched lymph node metastases.

E, Spatial mapping of Visium HD, colored by cell subtype.

F-I, Patient-level comparison of metastasis-associated cell populations between LN- and LN+ PeCa patients. Box plots show the patient-level abundance of S07_myCAF_ACTA2, CD4T06_Treg_FOXP3, CD8T05_Tc_ISG15, and CD4T08_Th_ISG15, respectively. Each point represents one patient.

J-L, Patient-level correlations between S07_myCAF_ACTA2 abundance and selected T-cell populations. Scatter plots show the associations of S07_myCAF_ACTA2 abundance with CD4T06_Treg_FOXP3, CD8T05_Tc_ISG15, and CD4T08_Th_ISG15, respectively. Correlation coefficients and corresponding p values are shown in the plots.

M-Q, Marker-gene expression used for manual validation of major cell-subtype annotations. Dot plots show representative marker genes for stromal cell subtypes (M), B-cell/plasma-cell subtypes (N), endothelial cell subtypes (O), myeloid cell subtypes (P), and T/ILC/NK cell subtypes (Q). Dot size represents the fraction of cells expressing each gene, and color intensity represents the mean expression level within each subtype.

R, Dot plot showing the expression of selected MHC class I and class II genes across all malignant epithelial subtypes. Dot size represents the fraction of cells expressing each gene, and color intensity represents the mean expression level within each subtype.

Figure S2 | Cancer cell subtype information.

A, The bar chart shows the proportion of each tumor-cell subtype among the total tumor-cell population.

B, The bar chart illustrates the changes in proportions of tumor subtypes relative to total tumor cells between the LN+ and LN- groups. Length represents the magnitude of change.

C, Violin plots illustrate the expression patterns of specific MHC class I and class II genes across C5-C8 subtypes.

D, Patient-level expression of metastasis-associated transcription factors in LN- and LN+ PeCa patients. Box plots show the expression levels of SOX2, ERG, HIF1A, FOS, FOXQ1, and TCF4 between the two groups. Each point represents one patient.

E, Patient-level SCENIC regulon activity of the identified transcription factors in LN- and LN+ PeCa patients. Box plots show differences in regulon activity between the two groups.

F-G, Functional enrichment analysis of SCENIC-predicted target genes of the identified transcription factors. Bubble plots show representative Gene Ontology biological processes (F) and KEGG pathways (G) associated with the predicted regulon targets.

H, Overlap between SCENIC-predicted transcription factor targets and genes associated with metastasis-related tumor programs. The heatmap summarizes the overlap of regulon target genes with functional gene sets related to tumor-cell states.

I, Associations between transcription factor regulon activity and metastasis-associated tumor programs across patients. The heatmap shows the relationships between regulon activity and tumor functional programs, including stemness, hypoxia-related, and immune-related signatures.

Figure S3 | Cell-cell communication information in PeCa.

A, The box plot shows cell-cell interaction strength between the LN+ and LN- groups.

B-D, Bar plot shows changes in signaling pathway activity associated with CD4 T cells, CD8 T cells and epithelial cells between the LN+ and LN- groups.

E, UMAP visualization of CAF populations included in the pseudotime analysis. Cells are colored according to CAF subtype, including resident fibroblast-like cells, S06_iCAF_IL6, and S07_myCAF_ACTA2.

F, Monocle3 pseudotime trajectory of CAF populations. The inferred trajectory illustrates continuous transcriptional-state transitions among resident fibroblast-like cells, iCAFs, and myCAF_ACTA2 cells.

G, Distribution of pseudotime values across CAF subtypes at the patient level. Violin and box plots show the pseudotime distribution of resident fibroblast-like cells, iCAFs, and myCAF_ACTA2 cells.

H, Changes in CAF-state composition along pseudotime. The stacked area plot shows the relative contribution of the three CAF states across scaled pseudotime.

I, Dynamic changes in representative functional programs along CAF pseudotime. Lines show changes in selected biological programs associated with fibroblast remodeling during pseudotime progression.

J, Dynamic expression of representative CAF-associated genes along pseudotime. The heatmap illustrates changes in genes associated with resident fibroblast, myofibroblastic, and extracellular-matrix-remodeling states.

Figure S4 | Validation of BHLHE41 knockdown and overexpression efficiency in CAFs by RT-qPCR.

A, Relative BHLHE41 mRNA expression in control CAFs, BHLHE41-overexpressing negative control CAFs, and BHLHE41-overexpressing CAFs.

B, Relative BHLHE41 mRNA expression in control CAFs, si-BHLHE41 negative control CAFs, and si-BHLHE41 CAFs.

C, Patient-level contribution of different CAF subpopulations to total CAF-derived POSTN transcripts. Each point represents one patient, and paired lines connect CAF subpopulations from the same patient. The y-axis indicates the fraction of POSTN UMIs contributed by each CAF subtype among total CAF-derived POSTN transcripts. Box plots summarize the distribution across patients for S06_iCAF, S07_myCAF, and S08_apCAF.

D, Spatial distribution of tumor cells in the corresponding tissue sections. Tumor cells are highlighted in yellow, whereas the remaining spatial locations are shown in gray.

E-G, Comparison of POSTN, ITGAV, and ITGB5 expression between the metastatic and non-metastatic groups.

H, Multiplex immunofluorescence staining of BHLHE41, POSTN, and ACTA2 in two independent PeCa cases. Representative images show DAPI (blue), ACTA2 (green), BHLHE41 (yellow), POSTN (red), and merged fluorescence signals.

I-J, Quantitative spatial association between POSTN and ITGAV/ITGB5 in the lymph-node section. (I) Same-spot and six-neighbor Spearman correlations between POSTN and ITGAV or ITGB5. (J) POSTN-high spots were not significantly closer to ITGAV- or ITGB5-high spots than expected from 999 random spot sets.

K-L, Quantitative spatial association between BHLHE41 and POSTN in the lymph-node section. (K) Same-spot and six-neighbor Spearman correlations. (L) High-expression BHLHE41 spots were significantly closer to POSTN-high spots than expected from 999 random spot sets.

Supplementary Table

Supplementary Table S1. Patient-level quality-control metrics and major cell-type composition of the penile cancer single-cell RNA-sequencing cohort.

Supplementary Table S2. Differentially expressed genes in malignant epithelial cells between LN+ and LN− penile cancer patients.

Supplementary Table S3. SCENIC-based analysis of transcription factor regulon activity.