Transcriptomic Profiling Reveals Cancer-Associated Fibroblasts as Potential Targets for the Prognosis and Treatment of Cervical Squamous Cell Carcinoma

Running title: CAFs in CSCC

Zhihua Ou,[1,3,13] Shitong Lin,[2,4,13] Jiaying Qiu,[1,5,13] Wencheng Ding,[2,4,13] Peidi Ren,[1,3,13] Dongsheng Chen,[1,13] Jiaxuan Wang,[1,3,13] Yihan Tong,[1,3,6] Di Wu,[1,3,7] Ao Chen,[1,8] Yuan Deng,[9] Mengnan Cheng,[1,5] Ting Peng,[2,4] Haorong Lu,[1,10,11] Huanming Yang,[1,12] Jian Wang,[1,12] Xin Jin,[1] Ding Ma,[2,4] Xun Xu,[1,11,*] Yanzhou Wang,[9,*] Junhua Li,[1,3,*] Peng Wu[2,4,*]

[1] BGI-Shenzhen, Shenzhen 518083, China.

[2] Cancer Biology Research Center (Key Laboratory of the Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430000, China.

[3] Shenzhen Key Laboratory of Unknown Pathogen Identification, BGI-Shenzhen, Shenzhen 518083, China.

[4] Department of Gynecologic Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430000, China.

[5] College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China.

[6] College of Innovation and Experiment, Northwest A&F University, Yangling 712100, China.

[7] School of Basic Medicine, Qingdao University, Qingdao 266071, China.

[8] Department of Biology, University of Copenhagen, Copenhagen DK-2200, Denmark.

[9] Department of Obstetrics and Gynecology, Southwest Hospital, Third Military Medical University, Chongqing 400038, China.

[10] China National GeneBank, BGI-Shenzhen, Shenzhen 518120, China.

[11] Guangdong Provincial Key Laboratory of Genome Read and Write, Shenzhen 518120, China.

[12] James D. Watson Institute of Genome Sciences, Hangzhou 310058, China.

[13] These authors contributed equally

[*] Correspondence: pengwu8626@tjh.tjmu.edu.cn (P.W.), lijunhua@genomics.cn (J.L.), w.y.z@foxmail.com (Y.W.), xuxun@genomics.cn (X.X.)

ABSTRACT

To understand the etiological, structural, and immunological characteristics of cervical squamous cell carcinoma (CSCC), we conducted single nucleus RNA sequencing (snRNA-seq) and spatial transcriptomics (ST) experiments for cervical samples from 20 individuals. When exploring the possible factors shaping the intra-individual immune heterogeneity in CSCC, we identified a cluster of cancer-associated fibroblasts (CAFs) enriched around some tumors, which highly expressed *ACTA2*, *POSTN, ITGB4,* and *FAP*. Results showed that the CAFs might support the growth and metastasis of tumors by inhibiting lymphocyte infiltration and remodeling the tumor extracellular matrix. Moreover, high CAF signals predicted poorer clinical outcomes for CSCC patients. Our data also revealed the infection profiles of HPV in tumors, the critical factors involved in the progression of cervical cancerous lesions, and the association between tumor metabolism and immune response intensity. Collectively, our findings may improve the prognostic and therapeutic methods for CSCC.

KEYWORDS

Cervical cancer; tumor microenvironment; cancer-associated fibroblasts; spatial transcriptomics; single-cell sequencing

INTRODUCTION

Cervical cancer is the fourth most common cancer affecting women's health globally, especially in low- and middle-income regions[1,2]. Since 2018, the World Health Organization (WHO) has called for the global elimination of cervical cancer, quantifying actions in vaccination, screening, and disease treatment/management[3], which require joint efforts from different parties for decades.

Currently, over 12 types of human papillomaviruses (HPVs) are carcinogenic[4]. Among them, HPV16 is responsible for 60%-70% of the cervical cancer cases, especially cervical squamous cell carcinoma (CSCC). The HPVs encode three oncoproteins, including E5, E6, and E7, which can disrupt the cell cycle and lead to uncontrolled cell division[5]. The development of CSCC generally goes through low-grade squamous intraepithelial lesion (LSIL), high-grade squamous intraepithelial lesion (HSIL), and invasive cervical cancer, which may need decades. The tumor

microenvironment (TME) of CSCC is formed by the complex interactions between viruses, cancer cells, stromal cells, and immunocytes, but was not fully characterized due to technical limitation and lack of study. Deciphering the factors propelling disease progression may give rise to novel prognosis and intervention strategies.

Although early cervical cancer receiving radical hysterectomy can achieve a favorable prognosis, the 5-year overall survival rate or disease-free survival rate of advanced cervical cancer are unsatisfactory[6,7]. At present, chemotherapy (e.g., paclitaxel, cisplatin, bevacizumab, etc.) and radiotherapy remain the main palliative treatment strategies for metastatic or recurrent patients, with a low response rate (48%) and a short survival period (17 months)[8–11]. Immunotherapy brings new hope to treating incurable cervical cancer by reversing the exhausted or suppressed immune activities. Immune-checkpoint blockade (ICB) drugs targeting programmed cell death 1 (PD1), programmed cell death ligand 1 (PD-L1), and cytotoxic T lymphocyte antigen 4 (CTLA4) are currently under trials for recurrent/metastatic cervical cancers[12,13]. Unfortunately, the overall response rates to ICB therapy were low, varying from 4% to 26%[12,14–16]. Clarifying the immune landscape of CSCC, especially the immunosuppression status in TME, may help us better address this phenomenon and adjust our treating strategy for cervical cancers.

Single-cell sequencing and spatial transcriptomics (ST) are state-of-the-art tools to unravel the cell heterogeneity and microenvironment of tumors, but applications of such techniques to CSCC investigation remain minimal. Only Hua et al. have reported the intra-tumoral heterogeneity of one cervical cancer patient using single-cell RNA sequencing in 2021[17]. In this study, we collected cervical samples from 20 individuals and combined single nucleus RNA sequencing (snRNA-seq) and ST to investigate the etiological, structural, and immunological profiles of CSCC. Deciphering the transcriptomic changes of HPV-induced CSCC would provide new insights into the diagnosis, prognosis, and treatment of CSCC, which may accelerate cervical cancer elimination.

RESULTS

snRNA-seq data revealed the cellular composition of CSCC

To fully characterize the cell composition of cervical tissues, we collected CSCC samples from 5 patients for snRNA-seq (**Figures 1A&S1, Table S1**). A total of 67,003 cells and 30,996 genes passed quality control (**Figure 1B, Table S2**), from which we identified 14 cell types based on canonical cell markers (**Figure 1B, Table S3**), including cancer cells (6,960), columnar epithelial cells (CECs, 22,396), endothelial cells (6,340), smooth muscle cells (4,502), fibroblasts (9,836), B cells (689), monocytes (5,281), T cells (4,930), regulatory T (Treg) cells (1,081), plasma cells (3,236), myeloid dendritic cells (DCs) (955), plasmacytoid DCs (272), mast cells (384) and natural killer (NK) cells (141). The uterine cervix contains two types of cells lining its surface, with stratified squamous epithelial cells on the ectocervix and simple columnar epithelial cells on the endocervix and crypts. The dysplasia of squamous epithelial cells leads to CSCC. Therefore, the cancer cells mainly expressed a known CSCC-associated gene *SERPINB3* (Serpin Family B Member 3)[18,19], tumor gene *TP63*, *CDKN2A,* and keratin gene *KRT15* of squamous cells (**Figure 1C&D**). Since the tissues were mainly from advanced cancer patients (FIGO Stage IB2-IIIC1; the tissue with no staging information was collected during cervical biopsy before chemotherapy), few normal epithelial squamous cells were isolated. Mapping of the snRNA-seq reads against high-risk HPV reference genomes revealed the presence of viral genes in cancer cells (**Figure 1C, Table S2**). We also identified a big cluster of columnar epithelial cells, which highly expressed *MUC5B* and *WFDC2*. This cell type is mainly located in the endocervix epithelia, but it can also appear at the squamocolumnar junction in the adult uterine cervix and some glands. Smooth muscle cells, fibroblasts, and endothelial cells are the major cell types composing the cervical stroma. The smooth muscle cells highly expressed *MYH11, MYLK*, *ACTG2, COL3A1, and COL1A1*, fibroblasts expressed high levels of *LAMA2* besides *COL3A1 and COL1A1*, while endothelial cells can be distinguished by high expressions of *EMCN, FLT1,* and *EGFLT* (**Figure 1C&D**). Besides the structural cells of the cervix, diverse immune cell types were also identified, with monocytes (*ITGAX, MX4A7*), T cells (*CD3E, CD247*), and plasma cells (*MZB1, IGKC*) being the most abundant (**Figure 1B&C**). In short, the snRNA-seq data provided a comprehensive atlas showing the structural and immune cell composition of the CSCC tissues.

Spatial transcriptomic characterization of CSCC

Spatial information is critical in understanding cell-cell interactions in tissues, which unfortunately was missing in the snRNA-seq data. Therefore, we utilized SpaTial Enhanced REsolution Omics-sequencing (Stereo-seq) technology[20] to acquire the *in situ* gene expression profiles. Cervical samples from 2 non-cancer and 14 CSCC patients were obtained and embedded in OCT (**Table S1**). Serial cryosections of 10 μm thickness were dissected from each OCT block for Stereo-seq, hematoxylin and eosin (HE) staining, and immunohistochemical (IHC) staining (**Figure S1A, Table S1**). Finally, a total of 18 ST slides (non-cancer, n=3; CSCC, n=15) were successfully obtained (**Figure 2A, Table S4**). The capture spots in Stereo-seq chips were 220 nm in diameter with a center-to-center distance of 500 nm between two adjacent spots. The capture spots were grouped into bins to include sufficient genes for accurate clustering. Our preliminary analysis revealed much higher RNA abundance in tumor areas than in stroma areas. To balance the expression differences between tumor and stroma, we annotated the CSCC ST slides at bin100 (100 x 100 spots) to fully demonstrate the tissue composition, which would cover an area of approximately 49.72 x 49.72 μm. The mean numbers of genes per bin for the CSCC ST slides ranged from 1,767 to 4,152 (**Table S4**). Because the three ST slides from non-cancer patients had lower gene expression intensity than CSCC slides, they were annotated at bin200 (99.72 x 99.72 μm). Uniform manifold approximation and projection (UMAP) analysis showed that bin clusters of CSCC and non-cancer tended to dissociate from each other while those of CSCC displayed some convergence (**Figure S1B**).

Considering the complex structure of cancerous tissues, we manually conducted the initial annotations of ST slides according to professional pathological assessments (based on HE and IHC staining results) and marker gene expression patterns (**Figure S1C&D)**. Six types of tissue clusters were generally identified in our CSCC samples, including tumor, stroma (without obvious inflammation), inflammation (stroma with diffuse inflammation or focal inflammation), gland, blood vessel, and necrosis. Tumor, stroma, and inflammation clusters were widely distributed among the ST slides of the CSCC samples, with certain samples containing necrosis, glands, and blood vessels (**Figure 2A**). Depending on the gene expression profiles, the tissue clusters may be further divided into sub-clusters with number suffixes (**Figure S1C**). It should be kept in mind that an ST cluster was virtually a mixture of

cells in a defined area but was designated by its major characteristics. For example, the tumor clusters were not purely composed of cancer cells, but may also contain other cell types though in small numbers, such as immune cells, fibroblasts, etc. Tissue-specific genes also displayed spatial patterns. Cancerous squamous cell-associated genes such *KRT5, CDKN2A,* and *SERPINB3* were mainly enriched in the ST tumor areas, *IGKC* and *IGLC2* enriched in inflammation areas, *VIM* in stromal areas, *ADRA2A* in blood vessels, and *MUC5B* in glands (**Figures S1D&S2**). In general, higher transcriptional and translational activities, cell proliferation, oxidative phosphorylation, and immune responses were observed in the ST tumor areas than in the other areas (**Figure 2B**). The initial manual annotations of the ST slides were used to assist downstream analysis of the TME.

Regarding the etiological role of HPV, the expression of viral genes in cervical tissues is always of interest. We mapped the ST sequencing reads against HPV reference genomes and found that all the 14 CSCC patients donating the 15 samples were HPV-positive, with 85.7% (12/14) infected by HPV16, 7.1% (1/14) by HPV33, and 7.1% (1/14) by HPV58. The HPV reads covered 8% - 100% (about 600bp - 7905bp) of the viral genome and were mainly identified in the ST tumor areas (**Table S4, Figures S3&S4**). Spatial visualization demonstrated varied capture signals of viral genes in different tumor sites of the same sample, with E5, E6, E7, and L1 frequently observed (**Figure S4**). In contrast, only marginal HPV reads were identified in the non-CSCC samples (**Table S4**).

ST data resolved the evolutionary trajectory of CSCC

Since ST data were able to reveal the spatial distribution of different tumor areas within the same tissue section, this offered us an opportunity to investigate the evolutionary trajectory of CSCC. High-grade intraepithelial neoplasia, carcinoma *in situ,* and invasive cancer represent different pathological statuses in the cervix, which frequently co-occur in the same patient. However, the dynamic changes of these pathological processes within the same individual remain obscure due to technical limitations. Here we first assessed the differentiation status of each tumor sub-clusters in the ST slides by calculating its CytoTRACE score[21]. HE staining was simultaneously used to identify cervical epithelia and their adjacent tumor areas. A cancerous epithelial region was deemed as a preinvasive cancerous lesion, while

the adjacent tumor area with a relatively higher CytoTRACE score, which indicated a relatively lower differentiation degree, was regarded as an invasive cancerous lesion. Because the metastatic path of tumors distant from the epithelia was hard to define, such tumor areas were not included for this analysis. Finally, we identified two samples (TJH37 and TJH90) from independent patients containing both preinvasive and invasive lesions (**Figure 2C&D**). Clusters 14 and 16 were recognized as preinvasive lesions, while clusters 12 and 13 were identified as the invasive lesions for TJH37 and TJH90, respectively.

Functional analysis based on the differentially expressed genes (DEGs) of preinvasive lesions versus invasive lesions ($Log_2FC > 0.5$ or $Log_2FC < -0.5$, $P < 0.001$) showed that the up-regulated genes mainly got involved in skin development, keratinocyte differentiation, response to wounding and bacterium, vitamin D receptor (VDR) pathway, blood vessel development and neutrophil degranulation, which were mainly associated with differentiation and intensified response to pernicious stimuli (e.g., HPV infection, epithelial injury). By contrast, the down-regulated genes were significantly correlated with RNA metabolism, oxidative phosphorylation, mitochondrion organization, cellular respiration, and viral mRNA translation, indicating increased metabolism in invasive cancer lesions (**Figure 2E**). The active engagement of the VDR pathway in the preinvasive lesion was worth noting, which suggested its essential role in the mucosal immunity of the cervix.

We then sought to pinpoint the key genes involved in the progression of cancerous lesions. A total of 16 DEGs ($Log_2FC > 1.0$ or $Log_2FC < -1.0$, $P < 0.001$) were shared by the two samples. Pseudo-time analysis showed that 12 genes (*ANXA1, AL136982.6, CRCT1, S100A7, S100A9, SPRR1A, SPRR1B, SPRR2A, SPRR2D, SPRR2E, KRTDAP,* and *KRT17*) were overexpressed by the preinvasive cancer lesions (**Figure 2F**), while the other 4 genes (*CD74, MT2A, KRT5,* and *KRT15*) tended to be overexpressed by the invasive lesions. Since an early prognosis is essential for the triage of patients with cervical intraepithelial neoplasia (CIN), we were curious if these DEGs could be of any clinical value. We downloaded the GSE63514 dataset from the Gene Expression Omnibus (GEO) database, which contained 66 cervical preinvasive (CIN) and 38 invasive lesion (cervical cancer) samples[22], and conducted PCA analysis to verify the discrimination ability of the

DEGs. Results showed that the combination of 11 DEGs (the rest 5 genes were not identified in this dataset) could well distinguish preinvasive (CIN) and invasive lesions (cervical cancer) (**Figure 2G)**. Random forest (RF) test revealed that these 11 genes (*ANXA1, S100A9, S100A7, KRT15, SPRR1A, SPRR1B, CD74, KRTDAP, MT2A, CRCT1,* and *KRT5*) had similar importance (**Figure 2H)**. The RF classification model also showed a good performance, reaching a mean AUC (area under the ROC curve) of 0.848 in 10 tests (**Figure 2I**). In summary, these DEGs may provide new insights into the differential diagnosis between CIN and cancer.

Variable immune inhibition in CSCC

Regarding the low response rates of ICB therapy to cervical cancer[11], we decided to scrutinize the immune landscape of CSCC for clues. The expression profiles of three gene sets with different immune functions, i.e., co-stimulatory, cytotoxic/effector, and co-inhibitory/exhaustion were evaluated in both snRNA-seq (**Figure 3A)** and ST data (**Figure 3B, Table S5**). At the single-cell level, the co-stimulatory genes were expressed in cells of both the innate and adaptive immunity, especially in Treg, T, and NK cells (**Figure 3A**). Treg cells were found to highly express *CD27, CD28, CD40LG, ICOS, TNFRSF18, TNFRSF4,* and *TNFRSF9.* While these genes are necessary for the maturation and normal suppressive function of Treg cells, overexpression of *CD27* in Treg cells may restrain the anti-tumor immune response[23,24]. Spatially, the co-stimulatory genes tended to be enriched in tumor and inflammation areas (**Figure 3B**). Especially, *TNFRSF18 (*also named glucocorticoid-induced TNF receptor*, GITR)* was commonly expressed in both inflammation and tumor areas. However, its expression was also up-regulated in the epithelia of non-cancer samples. In our snRNA-seq data, this gene was mainly detected in Treg, NK, T cells, and mast cells. Although *TNFRSF18* was associated with the immune suppression by Treg cells in tumors [25,26], its high spatial expression level may be jointly contributed by multiple types of immunocytes. The immune cytotoxic/effector genes were mainly expressed by T and NK cells, some of which were commonly up-regulated in the inflammation and tumor regions in the ST slides, including *GNLY, GZMA, GZMB,* and *NKG7* (**Figure 3B**). These genes were mainly expressed by NK cells (**Figure 3A**), indicating the indispensable role of NK cells in the cytotoxic response against CSCC. For the co-inhibitory/exhaustion genes, we failed to detect any prominent expression of *CTLA4* and *PD-1* in our ST data (**Figure 3B**), although

CTLA4 can be highly expressed by Treg cells and *PD-1* by Treg, T, and NK cells (**Figure 3A**). *PD-L1*, which was mainly expressed by plasmacytoid DCs, was only overexpressed in the tumor or inflammation areas of a small fraction of ST samples. While *CD276, ENTPD1, IDO1, LGALS9, and VSIR* were commonly detected in the ST samples, only *IDO1* and *LGALS9* seemed to have higher and wider expressions in the CSCC samples compared to the non-cancer samples (**Figure 3B**). These two genes were both expressed by DCs (**Figure 3A**) and could downregulate cytotoxic T cell activity[27,28]. Whether IDO1 and LGALS9 could be better targets for ICB therapy against CSCC than CTLA4 and PD-L1 remains to be explored. Moreover, when we zoomed in to check the immune genes in the same ST slide, their expressions can vary greatly between different tumor areas. In sample TJH08, the tumors commonly expressed high *IDO1*, low *PD-L1,* and very low *CTLA4* (**Figure 3C**). In contrast, only one tumor area in sample TJH37 expressed these genes (**Figures 3C&S2**). Collectively, although both our snRNA-seq and ST data displayed evidence of immune exhaustion in CSCC patients, the immune microenvironment varied considerably between and within patients.

Hypermetabolic tumors were associated with low immune response

Metabolism can modulate the immune microenvironment of tumors, which could be putative intervention targets for cancer therapy[29,30]. We performed gene set variation analysis (GSVA) separately on six pathways, including hypoxia, lactic acid, glycolysis, lipid metabolism, pentose phosphate, and oxidative phosphorylation pathways. The mean of the GSVA scores for the above six pathways was then calculated as the metabolic score of each tumor area. Based on the GSVA metabolic score, the ST tumor clusters ranked top 20 were categorized as hypermetabolic tumors, with those ranked the last 20 as hypometabolic tumors (**Figure 3D**). Generally, the hypermetabolic tumors displayed much higher activities in the oxidative phosphorylation, glycolysis, and lactic acid pathway, indicating active aerobic glycolysis in proliferating cancer cells, i.e., the Warburg effect. Moreover, the hypermetabolic tumors were also accompanied by severe hypoxia and active lipid metabolism, suggesting intense oxidative and nutrient stress in fast-growing tumors. A negative correlation between metabolism and immune response was further observed. We found that the hypermetabolic tumors experienced lower levels of adaptive and innate immune interference, lymphocyte infiltration, and

lymphangiogenesis **(Figure 3E**). We further explored the relationships between tumor differentiation, hypoxia, and immunity based on the gene module expression score in every bin (100 x 100 spots). A lower tumor differentiation level (indicated by a higher CytoTRACE score) was positively linked with severer hypoxia (r = 0.28, $P < 0.001$) (**Figure S5**). Furthermore, hypoxia was negatively associated with adaptive immunity (r = -0.19, $P < 0.001$), especially the infiltration of lymphocytes (r = -0.22, $P < 0.001$). Remarkably, the abundances of Th1 cell (r = -0.17, $P < 0.001$), B cell (r = -0.30, $P < 0.001$), immature B cell (r = -0.18, $P < 0.001$), and mast cell (r = -0.22, $P < 0.001$) were significantly lower in tumors with hypoxia. Although hypoxia did not significantly affect the innate immunity score, increased $CD56^-$ NK cell (r = 0.27, $P < 0.001$) and immature dendritic cell (r = 0.23, $P < 0.001$) were identified in tumors with hypoxia (**Figure S5**), which may not effectively combat cancer cells. We further tried to spatially verify the correlation between metabolism and immunity using *TOP2A* and *PTPRC* as markers for proliferating cells and immunocytes, respectively. In samples TJH34 and TJH35, we were able to identify both hyper- and hypo-metabolic tumor areas in the same ST slides. Much higher *PTPRC* expression was detected within and outside the hypometabolic tumor areas, consistent with the lymphocyte distribution pattern in HE images (**Figure 3F**). These observations suggested that the oxygen and nutrient deficit environment caused by actively proliferating tumors may prevent the accumulation of B cells, mast cells as well as functional DCs, NK cells, and helper T cells. Taken as a whole, poorly differentiated tumors tended to be more metabolically active, which may result in hypoxia and the paucity of immune intervention.

Identification of a cluster of cancer-associated fibroblasts (CAFs) in CSCC

When exploring the immune differences between the hyper- and hypo-metabolic tumor areas in sample TJH34, we noticed a unique spatial cluster outside the hypermetabolic tumor regions. This cluster was different from most stromal clusters and looked like a ribbon enclosing the tumor. We reckoned that this particular cluster might be critical in shaping the immune heterogeneity between tumor areas. As this cluster was part of the stroma, we closely scrutinized the fibroblasts in our snRNA-seq data. Luckily, we identified a small set of fibroblasts derived from all five samples (**Figure 4A**), which highly expressed reported marker genes for CAFs, including *ACTA2*, *POSTN, ITGB4,* and *FAP* (**Figure 4B**). The CAFs had a lower stemness

score than cancer cells and fibroblasts and were in various cell cycle stages (**Figure 4C**). Function enrichments based on the hallmark gene sets (MSigDB v7.4, https://www.gsea-msigdb.org/gsea/msigdb/) showed that CAFs shared common activities with both fibroblasts and cancer cells (**Figure 4D**). CAFs were involved in similar pathways to fibroblasts including UV response down, angiogenesis, myogenesis, and epithelial-mesenchymal transition. For pathways including the p53 pathway, KRAS signaling down, estrogen response, mitotic spindle, G2/M checkpoint, and E2F targets, CAFs showed similar activities to cancer cells. At the gene level, CAFs not only highly expressed marker genes for fibroblasts, such as the collagen protein family (*COL1A1, COL3A1, COL4A1, COL5A2, COL6A3*, etc.), but also marker genes for malignant squamous cells, such as *KRT4* and *KRT13* of the keratin family (**Figure 4E, Table S6**). Though CAFs in cancers might have different origins, the epithelial characteristics indicated that the origin of the CAFs in our CSCC samples might be associated with EMT.

To locate the spatial distribution of CAFs in CSCC tissues, we adopted the multimodal intersection analysis (MIA) approach developed by Moncada et al. to integrate snRNA-seq and ST data[31]. Briefly, this method calculated the overlapping degree of the expression levels of cell type-specific genes identified by snRNA-seq data and the area-specific genes characterized by ST data. The smaller the resultant p-value, which was mentioned as MIA score in our later description, the stronger the correlation between a defined cell type and an ST area. Initial MIA results showed that our ST clustering results complied with the expected cell composition in the corresponding areas (**Figure S6A**). Unfortunately, the MIA score alone cannot fully reflect the spatial specificity of cells, especially in areas with low RNA abundance. Therefore, a high expression level of *POSTN*, which was experimentally verified to be linked with CAFs [32–34], and a high MIA score for CAFs were simultaneously utilized to define ST clusters of CAFs (**Figure S6B**). Results showed that CAFs were enriched around some tumor areas in ST slides (**Figure 4F**), including the hypermetabolic tumor areas of sample TJH34 (**Figures 3F&4F**). The existence of CAFs in CSCC was further confirmed by IHC staining of POSTN using serial tissue sections of the same samples (**Figure S6C**). Notably, not all the tumor areas were surrounded by CAFs, making us curious about the biological differences associated with the presence of these cells.

CAFs might facilitate the growth and metastasis of CSCC from diverse aspects

To comprehensively reveal the biological functions of CAFs in CSCC, we divided the tumor areas in the ST slides into two types: tumor areas surrounded by CAFs (CAFs+ tumors), and tumor areas not surrounded by CAFs (CAFs- tumors). Three ST slides were found to contain both CAFs+ and CAFs- tumor areas and were used for downstream analysis. We then used the up- and down-regulated DEGs between the CAFs+ and CAFs- tumor areas of these 3 samples to conduct GO enrichment analysis (**Figures 4G&H, Table S7**). Results showed that the CAFs+ tumors were more active in energy usage, metabolism, mitosis, and cell growth than CAFs- tumors (**Figure 4I**). Meanwhile, cellular adhesion, apoptosis, and immune response were down-regulated in CAFs+ tumors. The above observations well coincided with the immune and metabolic heterogeneity of CSCC (**Figure 3D&E**), especially in sample TJH34 (**Figure 3F**). These indicated that the presence of CAFs can support tumor progression from different aspects.

Next, we calculated the gene module expression scores of 993 individual cells regarding immune gene sets to evaluate the immune cell abundances. Results showed significantly reduced numbers of B cells, CD4 T cells, CD8 T cells, neutrophils, DCs, NK cells, and Th1 cells in CAFs+ tumors (**Figures 5A&S6D**). Intriguingly, more macrophages, possibly tumor-associated macrophages (TAM), were identified in CAFs+ tumors (**Figures 5A&S6E**), which may promote the proliferation and migration of cancer cells [35]. Therefore, CAFs not only acted as a physical barrier to prevent the infiltration of pro-immunity cells into tumor areas but might also recruit anti-immunity macrophages to facilitate the growth of tumors.

As part of the stroma, CAFs will have to interact closely with cancer cells, stromal cells, and immune cells. Indeed, analysis of the snRNA-seq data showed complicated interactions between CAFs and the other cells regarding extracellular matrix (ECM) formation and cell-cell contact (**Figure 5B**). CAFs highly expressed genes of the collagen family, especially *COL1A1, COL1A2, COL4A1, COL4A2, COL4A5, COL6A1, COL6A2,* and *COL6A3*) to interact with *CD44* expressed by immunocytes and smooth muscle cells, which may be involved in cell adhesion and migration. The collagens also interact with diverse members of the integrin family

expressed by the cancer cells, immune cells, and stromal cells. Similarly, *FN1* (fibronectin 1, a soluble glycoprotein) and laminins (*LAMA2, LAMA3, LAMA4, LAMA5, LAMB1, LAMB2, LAMB3, LAMC1, LAMC3*) expressed by CAFs also interacted with the other cell types through integrins. The integrins are membrane receptor proteins made up of α and β subunits and are involved in cell adhesion and recognition. Interestingly, CAFs used different heterodimeric forms of integrins to contact the other cell types. They may interact with cancer cells through integrins composed of subunits α2β1, α3β1, and αvβ8, while interacting with endothelial cells, CECs, smooth muscle cells through integrins made up of subunits α9β1, α6β1, and α1β1, and with T cells and NK cells through integrins made up by subunits α1β1 (**Figure 5C**). Importantly, CAFs may take advantage of F11R (also called JAM1, junctional adhesion molecule 1) to form tight junctions with cancer cells and stromal cells through F11R and JAM3, which might prevent the infiltration of immunocytes. CAFs might also express other matrix proteins including *THBS1* (thrombospondin 1), *THBS2,* and *TNC* (tenascin C) to communicate with immunocytes, smooth muscle cells, cancer cells, and plasma cells through *CD44*, integrin (α3β1), and *SDC1*. Moreover, CAFs overexpressed several tissue remodeling factors (**Figure 5D**), including *POSTN* (periostin, a secreted extracellular matrix protein), *FAP* (fibroblast activation protein, a serine protease), *MMP1* (matrix metalloproteinase 1), *TNC* (Tenascin-C, a matrix protein), and *LOXL1* (lysyl oxidase like 1, catalyzes the cross-linking of collagen and elastin). This evidence suggested a critical role of CAFs in shaping the tumor extracellular environment.

Besides their role in ECM construction, CAFs may also enhance the stemness and proliferation of cancer cells through overexpressing secreted factors including *SEMA3C, POSTN, CXCL6* (**Figure 5B**). *SEMA3C* was reported to promote cancer stem cell maintenance, angiogenesis, and invasion[36,37]. *POSTN* can augment cancer cell survival by activating the Akt/PKB pathway through integrins αvβ3[38]. It may also promote cancer growth through the PTK7-Wnt/β-Catenin signally pathway[33]. *CXCL6* (C-X-C motif chemokine ligand 6), though mainly related to immune response, was reported to promote the growth and metastasis of esophageal squamous cell carcinoma[39] (**Figure 5D**). Another highly expressed gene in CAFs, *SNAI2* (Slug), a snail-related zinc-finger transcription factor, may inhibit apoptosis and promote cancer progression[40,41]. Several common growth factors such as *TGFB1*

(transforming growth factor beta 1), *EGF* (epidermal growth factor), and *VEGFA* (vascular endothelial growth factor A) were also expressed by CAFs (**Figure 5D**). Moreover, the upregulation of *LSD1* (histone lysine demethylase 1) in CAFs might inhibit the IFN activation to evade immune attack[42]. The Wnt5a signaling protein produced by CAFs might also suppress the immune response to facilitate tumor metastasis (**Figure 5B**)[43,44]. In summary, CAFs might be able to potentiate the TME, promoting the progression of tumors.

The presence of CAFs was associated with poorer outcomes of CSCC patients

To verify the pro-tumorigenic effects of CAFs, we first performed survival analyses with a dataset from the Cancer Genome Atlas (TCGA)[45], which contained 252 CSCC patients. The GSVA score for CAFs was calculated for each CSCC patient using the marker gene set (*ACTA2*, *POSTN, ITGB4,* and *FAP*) (**Figure 4B**). Not surprisingly, higher signals of CAFs predicted unfavorable progression-free survival (HR = 1.66, 95%CI = 1.03-2.67, P = 0.038) and overall survival (HR = 1.69, 95%CI= 1.00-2.84, P = 0.05) for CSCC patients (**Figure 5E**). Next, we measured the expression levels of POSTN, a biomarker of CAFs, in the stroma adjacent to cancer lesions of 61 archived formalin-fixed paraffin-embedded (FFPE) CSCC samples (**Figure 5F**). Correlation analysis revealed that higher expression levels of POSTN were associated with poorer differentiation, more advanced tumor stages, more frequent lymph node metastasis, and higher squamous cell antigen (SCC) concentrations in peripheral blood (**Figure 5G**), which further confirmed the pro-tumorigenic ability of CAFs.

Collectively, our results indicated CAFs were a crucial component of the TME of CSCC, forming a barrier to protect the cancer cells from immune surveillance and clearance, secreting cytokines to stimulate cell proliferation and angiogenesis, inhibiting apoptosis, and remodeling the ECM to enhance tumor metastasis (**Figure 5H**). Considering the deteriorating effect of CAFs, they may serve as potential targets for the prognosis and treatment of CSCC.

DISCUSSION

Although vaccines and radical hysterectomy are effective measures in preventing and treating cervical cancers, the treatment of recurrent/metastatic cervical cancers

remains a big obstacle to achieving the goal of cervical cancer elimination. Herein, we have characterized a high-resolution transcriptomic landscape of CSCC combining snRNA-seq and ST technology, providing detailed descriptions of the etiological, structural, and immunological characteristics of CSCC, which may facilitate the management and treatment of HPV-induced cervical cancer.

Our data showed that HPVs were active in transcription or translation in cancerous cervical squamous cells, with oncogenic genes (E5, E6, and E7) highly expressed. Because HPVs have DNA genomes, the recovery of almost intact viral genomes from transcriptomic data was vivid evidence suggesting active viral activities in advanced tumors. The HPV genomes contain two polyadenylation sites[46]. An early polyadenylation (A_E) site locates between E5 and L2, which is shared by transcripts of E6, E7, E1, E2, E4, and E5. A late polyadenylation (A_L) site locates after the L1 open reading frame, which is utilized by transcripts of L1, L2, and even E7 and E4. Since Stereo-seq used polyT to capture mRNA, the genes adjacent to the polyadenylation sites may have a higher chance to be sequenced, which was reflected by the general higher expressions of E5 and L1 than the other genes (**Figure S3**). Still, the early-stage genes, especially E5, E6, and E7, tended to have higher expressions than the late-stage genes including L1 and L2. Because HPVs assemble and release progeny virion in differentiated squamous cells, higher expressions of early-stage genes than late-stage genes in proliferating cancer cells were reasonable. Moreover, though not adjacent to the A_E site, oncogenes E6 and E7 displayed similar or even higher expression levels than E5 in tumors (**Figures S3&S4**), indicating their pivotal role in the tumorigenesis of cervical squamous cells. Indeed, therapeutic vaccines targeting viral E6 and E7 genes have shown some promising results in treating cervical precancers and cancers[47–49].

Traditional bulk-RNA sequencing and single-cell RNA sequencing are generally insufficient to obtain gene expression profiles of consecutive pathological statuses. Herein, we revealed key genes and pathways associated with dysplasia progression with cutting-edge ST methods using tissue samples containing both preinvasive and invasive cancer lesions. The VDR pathway up-regulated in the preinvasive status (**Figure 2E**) was found to be critical in maintaining mucosal barrier homeostasis and preventing microbial infections[50–52]. The VDR pathway can be up-regulated by Toll-

like receptors (TLRs) to induce antimicrobial factors such as cathelicidin in humans[53]. Moreover, the administration of vitamin D was reported to improve pulmonary lesions caused by tuberculosis infection[54]. Therefore, the TLR-VDR axis might be pivotal in combating HPV infection. Agonists of the TLR family, such as imidazoquinolines and resiquimod, and vitamin D, might be possible therapeutics to manage precancer and cancer lesions[55,56]. Besides, the 16 DEGs between preinvasive and invasive lesions (**Figure 2F**) demonstrated good performance in discriminating CIN from cervical cancer in our test data. The four up-regulated genes, including *CD74, MT2A, KRT5,* and *KRT15*, were associated with tumor progression in several types of squamous cell carcinomas[57–60], their prognostic values in CSCC may worth further investigations.

Nowadays, ICB therapies, especially those using PD-L1/PD-1 and CTLA4 inhibitors, are among the novel methods to treat metastatic cervical cancers. Several studies reported wide expression of PD-L1 in cervical cancers, with positive rates ranging from 34% to 96%[61–63]. However, PD-L1 expression alone wasn't associated with the disease outcome of cervical cancer patients[61]. Indeed, the response rates to PD-1/PD-L1 and CTLA-4 inhibitors fluctuated greatly among different trials and the efficacies seemed independent of the expression status of associated checkpoint genes[11,64]. In our study, the expression levels of most immune checkpoint genes in tumor and inflammation areas of CSCC were not significantly higher than those in the non-cancer samples except for *LGALS9* and *IDO1* (**Figure 3B**). LGALS9 (i.e., galectin 9) downregulates effector T cell immunity through binding to Tim-3 on the T cell surface or inhibiting the antigen-presenting ability of DCs[27,28]. While disruption of the galectin 9 signaling pathway was shown to induce tumor regression in mice harboring pancreatic ductal adenocarcinoma, reversed effect was reported in lung metastasis mouse models[65,66]. IDO1 is mainly expressed in DCs and helps degrade tryptophan into kynurenine, which suppresses T cell functions[67]. It is found that inhibition of IDO1 could enhance the radiosensitivity of HeLa and SiHa tumorsphere cells, indicating the potential application of IDO1 inhibitors as radiosensitizers[68]. Whether targeting LGALS9 and IDO1 can improve the treatment of CSCC requires further exploration. Our study also revealed insufficient immune surveillance in tumor areas with high oxygen and nutrient stress, highlighting the critical role of metabolic modulation in TME. Indeed, recent clinical trials have combined checkpoint inhibitors

with metabolic agents targeting glucose, amino acid, and nucleotide metabolism[69]. A better understanding of the crosstalk between immune response and metabolism would further benefit cancer therapies.

CAFs have been characterized in multiple types of cancers and can be classified into diverse subtypes[70]. While cell line studies have implicated a supportive role of CAFs in the proliferation of cervical cancer cells[71,72], this is the first study to systemically describe the spatial distribution and biological properties of pro-tumorigenic CAFs in clinical samples of CSCC. Due to tissue heterogeneity, the marker genes for CAFs varied among cancers. In this study, *ACTA2*, *POSTN, ITGB4,* and *FAP* were adequate marker genes to identify CAFs in CSCC (**Figure 4B**). Other genes such as *KRT4, ITGA1, COL24A1,* and *COL7A1* might serve as complementary marker genes for CAFs in CSCC (**Figure 4E, Table S6**), which displayed cellular properties of both fibroblasts and cancerous squamous cells. These CAFs contributed significantly to the heterogeneity of TME, which displayed a pro-tumorigenic phenotype by facilitating tumor growth, metastasis, and immune evasion. The CAFs+ tumors were active in proliferation but lacked lymphocyte infiltration. Exposing these immune-evasive tumors to the immune system is essential to eradicate the cancer cells. Therefore, it is recommendable to incorporate therapies targeting CAFs when treating patients with advanced CSCC. Researchers have tried to interfere with the activation, the action, and the normalization processes of CAFs using antibodies or inhibitors, with several clinical trials ongoing[73]. Since genes highly expressed by CAFs are also essential to normal tissues, their efficacies and side effects require close monitoring.

In conclusion, our data systemically demonstrates the high heterogeneity of viral gene expression, immune response, and metabolism in CSCC, indicating that combined drugs or therapies targeting multiple biological processes would be better practice to treat CSCC. Besides therapeutic vaccines and ICB therapies, our findings suggest that interventions on CAFs and tumor metabolism may complement the current treatments of CSCC. Further investigations into these biological aspects may facilitate the development of new drugs or therapies against CSCC and the other HPV-induced squamous cell carcinomas.

METHODS AND MATERIALS

Patients and samples

The cervical specimens were collected from 20 patients aged 38 to 69 by the Department of Obstetrics and Gynecology of Tongji Hospital in Wuhan and the Department of Obstetrics and Gynecology of Southwest Hospital in Chongqing. Based on colposcopy examination, 18 patients were diagnosed with CSCC (Stage IB1 to Stage IIIC1), 2 patients were diagnosed with benign gynecological diseases but also required surgery (**Table S1, Figure S1A**). Carcinoma staging was conducted based on the criteria of the FIGO staging system. The freshly-taken samples were used for single-cell RNA sequencing and ST experiments. A total of 61 archived FFPE samples were obtained to verify the presence of CAFs in CSCC.

Experiments

Single-cell sequencing

The collected cervical tissues were quick-frozen with liquid nitrogen for 30 minutes and then stored in a -80°C refrigerator. Nuclei isolation and permeabilization were performed under the guidance of Chromium Next GEM Single Cell Multiome ATAC + Gene Expression User Guide (CG000338). snRNA-seq libraries were prepared using the Chromium Single Cell 3 Reagent Kits v3 (10x Genomics, USA), according to the manufacturer's instructions. Briefly, high-quality sequencing data was obtained after a series of experimental procedures including cell counting and quality control, gel beads-in-emulsion (GEMs) generation and barcoding, post GEM-RT cleanup, cDNA amplification, gene expression library construction, and NovaSeq platform (Illumina, USA) sequencing.

Tissue preparation for spatial transcriptomic experiment

A tissue block with an edge length of less than 1cm was dissected from the surgically removed tissues. The tissue block was then rinsed by cold PBS, immersed in the pre-cooled tissue storage solution (Miltenyi Biotec, Germany), and then embedded with pre-cooled OCT (Sakura, USA) in a -30°C microtome (Thermo Fisher, USA) within 30 minutes after surgery. Three to four serial cryosections of 10 μm thickness were cut from the OCT-embedded samples for HE staining, Stereo-seq library preparation, and IHC staining. Brightfield images of the HE samples were

taken with a Motic microscope scanner (Motic, China) for histopathological assessment.

Quality control of RNA obtained from OCT-embedded samples

100-200 μm thick sections were cut from each OCT-embedded sample for total RNA extraction using the RNeasy Mini Kit (Qiagen, USA) according to the manufacturer's protocol. RNA integrity number (RIN) was determined by a 2100 Bioanalyzer (Agilent, USA). Only samples with RIN≥7 were qualified for the transcriptomic study. All samples used had RIN of 7-10.

Stereo-seq library preparation and sequencing

The spatial transcriptomic RNA library was constructed using Stereo-seq capture chips (BGI-Shenzhen, China), which had a size of 1 cm^2. The capture spots were 220 nm in diameter with a center-to-center distance of 500 nm between each other. Each Stereo-seq capture sequence contained a 25bp coordinate identity barcode, a 10bp molecular identity barcode, and a 22bp polyT tail for *in situ* mRNA hybridization[20]. A cryosection of 10 μm thickness cut from OCT-embedded tissue was quickly placed on the chip, incubated at 37°C for 3 minutes, and then fixed in pre-cooled methanol at -20°C for 40 minutes. The fixed tissue section was stained with the Qubit ssDNA dye (Thermo Fisher, USA) to check tissue integrity before fluorescent imaging. After that, the tissue section was permeabilized using 0.1% pepsin (Sigma, USA) in 0.01 N HCl buffer, incubated at 37°C for 14 minutes, and then washed with 0.1x SSC. RNA released from the permeabilized tissue was reverse transcribed for 1 hour at 42°C. Later, the tissue section was digested with tissue removal buffer at 42°C for 30 min. The cDNA-containing chip was then subjected to cDNA release enzyme treatment overnight at 55°C. The released cDNA was further amplified with cDNA HIFI PCR mix (MGI, China). Around 20ng cDNA was fragmented to 400-600bp, amplified for 13 cycles, and purified to generate DNA nanoball library, which was sequenced with the single-end 50+100bp strategy on an MGI DNBSEQ sequencer (MGI, China).

Immunohistochemical (IHC) staining

The IHC staining for Ki-67 and POSTN were performed under the manufactures' protocol. The frozen sections dried at room temperature were placed in a 37°C oven

for 10-20 minutes, fixed with 4% paraformaldehyde for 20 minutes, and washed thrice with PBS (pH = 7.4) for 5 minutes. The antigens were then repaired with EDTA (pH9.0) and the endogenous peroxidase was blocked by 3% hydrogen peroxide. The slides were further blocked with 3% BSA (G5001-100g, Servicebio) at room temperature for 30 minutes, and then incubated with Ki-67 (ab16667, Abcam, 1:200) or POSTN (ab215199, Abcam, 1:500) at 4°C overnight. Finally, the frozen slices were subjected to secondary antibody blocking, DAB staining, nuclear re-staining, and dehydration. The protein expression levels of Ki-67 and POSTN were evaluated by professional pathologists under a microscope. The expression scores of POSTN in FFPE samples were measured according to the positivity percentage (<5% = 0, 5-25% = 1, 26-50% = 2, 51-75% = 3, >75% = 4) and staining intensity (negative = 0, weak = 1, moderate = 2, strong =3) by two independent clinicians.

Bioinformatic analysis

Cell type characterization using snRNA-seq data

Quality control and gene expression quantification of snRNA-seq data

Raw sequencing files were first processed using CellRanger version v6.0.2 (10x Genomics, USA) to obtain gene expression matrices. After cell calling, the droplets containing no cell were excluded based on the number of filtered unique molecular identifiers (UMIs) mapped to each cell barcode. Droplets with low-quality cells or more than one cell were also removed. To obtain a gene expression matrix optimized to individual samples, the R package scCancer v2.2.1[74] was employed to further filter the expression matrix. The filtering thresholds were determined by catching outliers from the distribution of four quality spectra, including the number of total UMIs, the number of expressed genes, the percentages of UMIs from mitochondrial genes, and the percentages of UMIs from ribosomal genes. Besides filtering cells, genes expressed in less than three cells were also excluded to avoid false-positive results. The filtering thresholds for the five samples were documented in **Table S2**.

Cell type clustering using multi-sample snRNA-seq data

Integrative analysis of the snRNA-seq data from the five patients was carried out using the IntegrateData function of Seurat v4[75]. Further analysis included normalization, log-transformation, highly variable genes identification, dimension

reduction, clustering, and differential expression analysis were all conducted using default parameters of Seurat except that *dims* was set as 1:30. Initially, a total of 35 cell clusters were obtained (69,312 cells with 30,996 genes). To ensure reliable identification, we removed the cell clusters consisting of less than two samples and with less than 15 cells per sample. Finally, 14 cell clusters (67,003 cells with 30,996 genes) were determined based on reported cell type marker genes (**Table S3**).

Analysis of differentially expressed genes (DEGs)

Expression of each gene in each cluster was compared against the rest of the clusters using the Wilcoxon rank-sum test with the FindAllMarkers function of Seurat v4[75]. Significantly up- or down-regulated genes were identified using the following criteria: 1) the difference in gene expression level was >1.18 fold unless explicitly noted; 2) genes were expressed by more than 25% of the cells belonging to the target cluster. 3) the adjusted p-value was less than 0.05.

Processing and annotation of ST data

Preliminary processing of Stereo-seq data

Stereo-seq raw data were automatically processed using the BGI Stereomics analytical pipeline (http://stereomap.cngb.org/), where the reads were decoded, trimmed, deduplicated, and mapped against the human and HPV reference genomes. The reference genomes were: Human, GRCh38.p12; HPV16:K02718.1; HPV18: EF202147.1; HPV33: M12732.1; HPV58: D90400.1. Data of the chip area covered by tissue was extracted based on the ssDNA and HE staining images using the Lasso function of the BGI Stereomics website. It's worth noting that tumor sites usually had much higher overall mRNA levels than the other anatomical areas, leading to a significant imbalance of transcriptomic signals between the tumor areas and the other sites on the Stereo-seq slides. Therefore, to fully reflect the spatial transcriptomic landscape around the tumor areas, a bin size of 100 (100 spots x 100 spots, i.e., 49.72 x 49.72 μm) was used as the analytical unit for the annotation of CSCC ST slides, while a bin size of 200 (200 spots x 200 spots, i.e., 99.72 x 99.72 μm) was used for the non-CSCC samples. The downloaded data was then processed with Seurat v4[75]. We used the criteria of >200 UMIs per bin to remove bins with low expression signals. The data were then normalized using the SCTransform function. Dimension reduction was conducted with principal PCA.

Unsupervised clustering of bins was performed with UMAP. Sequencing and analytical details can be found in **Table S4**.

Annotation of bin clusters in ST slides

The bin clusters were annotated based on the *in situ* expression patterns of marker genes combining HE and IHC staining results. The spatial expression patterns of genes in ST slides (**Figures S1&S2**) were conducted with the SpatialFeaturePlot function of Seurat v4[75]. The HE and IHC images were examined by professional pathologists to determine the tissue types. The annotated ST areas were confirmed to be consistent with the HE and IHC assessment and marker gene expression patterns.

Identification of viral RNA

The viral reads were mapped against HPV reference genomes with BWA. The genome coverage (covered length/full length of the reference genome) and effective depth (total mapped bases/covered length) for each type were calculated. Only samples with a viral genome coverage of >5% and an effective depth of >50x were deemed as HPV positive (**Table S2&S4, Figure S3**).

GO enrichment analysis

All the GO enrichment analysis in this study was conducted using Metascape[76] with default settings (**Figures 2B, 2E&4I**). The reference species was *H. sapiens*.

Analysis of the progression trajectory of cervical cancerous lesions with ST data

CytoTRACE[21] was used to estimate the relative differentiation status of sub-clusters of the tumor areas in ST slides with default settings. The identification of DEGs and GO enrichment analysis were conducted as mentioned above. According to the annotation clusters of each ST slide, pseudotime analysis of preinvasive and invasive cancerous lesions was conducted with Monocle3 v1.0.0[77]. To test the discriminative ability of the DEGs, we performed principal components analysis (PCA) in R with 104 CIN/cancer samples from the GSE63514 dataset downloaded from the Gene Expression Omnibus (GEO) database. To access the importance of each gene in discriminating CIN and cancer, we implemented a random forest analysis using

the randomForest (v4.6-14) package in R, which used the decrease of Gini index impurity as a splitting criterion. We ran ten random forests consisting of 1000 trees to assess the partitioning ability for cancer and CIN for each gene. In each RF test, we randomly selected 70% of the CIN and cancer samples from the PCA dataset to train the random forest model. The remaining 30% samples were used to verify the effectiveness of the RF model with the pROC R package, implementing 2000 bootstrap replications.

Characterization of the immunity, hypoxia, and metabolism statuses in CSCC

Signature enrichment analysis of ST clusters

In enrichment analysis, the expression scores of signature genes were calculated for individual bins using the AddModuleScore (on log-normalized data) of Seurat v4[75] with default parameters. Pathways and cell types included in the enrichment analysis, with the corresponding reference for gene signatures, were listed as follows: hypoxia[78], angiogenesis (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/), cytolytic activity[79], lymphangiogenesis[80], glycolysis[81], lipid metabolism[82], pentose phosphate pathway[83], lymphocytes[84], macrophages[85], and other immune cells[86].

Prediction of the spatial distribution of immunocytes

Using the gene signatures of immunocytes as input, we calculated the cell type scores of each bin in tumor areas using the AddModuleScore of Seurat v4[75]. Combining with the spatial coordinate of bins, the possible spatial distribution of the corresponding cell type was obtained (**Figure S6D**).

Characterization of CAFs

Signature enrichment analysis of fibroblasts, CAFs, and cancer cells with snRNA-seq data

Expression scores of signature genes from MSigDB v7.4 (https://www.gsea-msigdb.org/gsea/msigdb/) were calculated for individual cells using the AddModuleScore function (on log-normalized data) of Seurat v4[75] with default parameters to assess differential pathways in fibroblasts, CAFs, and cancer cells.

Cell cycle and stemness analysis of fibroblasts, CAFs, and cancer cells with snRNA-seq data

We used the AddModuleScore function of Seurat v4 to calculate the relative average expression of a list of G2/M and S phase markers to obtain the cell cycle scores[87]. Cell stemness analysis was conducted using the OCLR model and the stemness signatures embedded in the scCancer package[74,88].

Multimodal intersection analysis (MIA)

To integrate snRNA-seq and ST data, we calculated the overlapping degree of the expression levels of cell type-specific genes identified by snRNA-seq data and the area-specific genes characterized by ST data using the MIA approach[31]. The lower the p-value, the higher overlapping between a certain cell type and an ST area. MIA was conducted to confirm the consistency between cell types and ST annotated areas and to identify the ST areas composed of CAFs.

Cell-cell communication between CAFs and the other cell types in CSCC tissues

To understand the communication network between CAFs and the other cell types, we conducted cell-cell communication with CellChat[89] with the snRNA-seq data to obtain the ligand-receptor pairs regulated by CAFs.

Prognostic analysis of CAFs with TCGA data

The gene expression profiles of CSCC were downloaded from The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/) with the latest follow-up prognostic information obtained from an integrated clinical data resource. We calculated a signature score of CAFs for each CSCC patient with GSVA using the marker genes of CAFs (*ACTA2*, *POSTN, ITGB4,* and *FAP*). Based on the median GSVA score, the patients were then divided into two groups: high CAFs v.s. low CAFs. The Kaplan-Meier overall survival and progression-free survival curves were generated with GraphPad Prism 6.

Statistical analysis and plotting

Statistical analyses, including Student's t-test, Wilcoxon's rank-sum test and Wilcoxon signed-rank test, were performed in R 3.6.0. Asterisks indicate the significance levels of p-values: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. The schematic plots were created with BioRender (https://biorender.com/).

ETHICAL STATEMENT

This study was reviewed and approved by the Medical Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology (TJ-IRB20210609), Southwest Hospital, Third Military Medical University (KY2020142), and the Institutional Review Board of Beijing Genomics Institute, Shenzhen, China (BGI-IRB 21050).

AUTHOR CONTRIBUTIONS

Z.O., P.W., J.L., W.D., and D.C. designed this study. P.W., Z.O., J.L., X.X., D.M., X.J., H.Y., and J.W. supervised this study. P.W., Y.W., W.D., and S.L. coordinated sample collection. S.L., Y.D., and T.P. conducted snRNA-seq and IHC experiments. P.R., Y.T., and D.W. performed Stereo-seq library construction and H&E staining. A.C. and M.C. provided technical support for Stereo-seq experiments. H.R.L. conducted sequencing for Stereo-seq libraries. J.Q., J.W., Y.T., and D.W. conducted data analysis. Z.O., S.L, P.R., J.Q., and J.W. wrote the manuscript. Z.O., S.L., P.W., J.L., W.D., and D.C. reviewed and polished the manuscript.

DECLARATIONS OF INTEREST

A.C. and M.C. are applying for patents covering the chip, procedure, and applications of Stereo-seq. The other authors declare that they have no competing interests.

ACKNOWLEDGMENTS

This work received funding from the National Key R&D Program of China (2021YFC2701201), the National Natural Science Foundation of China (82072895, 82103134) and the Guangdong Provincial Key Laboratory of Genome Read and Write (2017B030301011). We thank China National GeneBank for providing sequencing services for this project. The authors also would like to thank Miss Feiyun Xuanyuan and Mr. Geer Xuanyuan for their inspirational communications.

DATA AND CODE AVAILABILITY

The data supporting the findings of this study will be deposited into CNSA (CNGB Sequence Archive) of CNGBdb (https://db.cngb.org/cnsa/) upon publication. All

analysis scripts are available on reasonable request. The dataset used to verify the distinguishing ability of the DEGs between preinvasive and invasive cancerous lesions was obtained from the Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo) by the Accession Number of GSE63514.

FIGURE LEGENDS

Figure 1. The single-cell atlas of CSCC tissues. (A) Workflow of snRNA-seq and Stereo-seq experiments applied to cervical tissues. n indicates the number of samples. (B) UMAP of cells identified from the snRNA-seq data of five CSCC tissues (left). The cell number and proportion of each cell type from each sample (right). (C) Expression of selected marker genes and HPV genes in the major cell types of CSCC tissues. (D) Expression matrix of cell-type marker genes in the 14 cell types isolated from CSCC tissues.

Figure 2. Progression trajectory analysis of CSCC. (A) Distribution of ST clusters in the 18 cervical samples. CSCC: tumor, stroma, inflammation, gland, blood vessel, and necrosis; non-cancer: epithelia and stroma. (B) Gene ontology (GO) enrichment of major ST clusters. (C) CytoTRACE scores of tumor sub-clusters in two ST slides. (D) HE images showing preinvasive cancerous lesions (blue) and invasive cancerous lesions (red) in two CSCC samples. (E) Dot plot showing GO terms enriched by the up- and down-regulated DEGs between preinvasive and invasive cancerous lesions. Preinvasive lesion was compared to invasive lesions. (F) Top, the clustering and pseudotime plots of the two ST samples. Bottom, heatmap showing the pseudotime expressions of the 16 DEGs shared by the two ST samples. (G) PCA plot showing the distinguishing ability of the 11 DEGs for cervical cancer and CIN (cervical intraepithelial neoplasia). (H) The decrease Gini scores for the 11 genes obtained in 10 random forest analyses. (I) AUC plot showing the classification accuracies of the 11 genes in 10 tests.

Figure 3. Transcriptomic analysis of the immune heterogeneity in CSCC. (A) Heatmap showing the expressions of gene sets associated with different immune functions (co-stimulatory, cytotoxic/effector, and co-inhibitory/exhaustion) in cell types identified by snRNA-seq data. (B) Heatmap showing the expressions of gene sets associated with different immune functions (co-stimulatory, cytotoxic/effector,

and co-inhibitory/exhaustion) in ST clusters. (C) Spatial expression of selected immune checkpoint genes (*PD-L1, CTLA4,* and *IDO1*) in two representative ST samples. The gene tended to be enriched in tumor areas. (D) Top, metabolic scores corresponding to the tumor areas of the 15 CSCC samples. Bottom, heatmap showing the GSVA scores of hypoxia, lactic acid, glycolysis, lipid metabolism, pentose phosphate, and oxidative phosphorylation pathways for tumor areas from each ST sample. (E) Violin plots showing differences in lymphocyte infiltration, adaptive immunity, innate immunity, and lymphangiogenesis between hyper- and hypo-metabolic tumors. The Y axis shows the GSVA scores for each pathway. The p-values were determined by the Wilcoxon signed-rank test. (F) Spatial expression of marker genes for proliferating cells (*TOP2A*) and immunocytes (*PTPRC*) in representative ST slides. The blue line indicates the border of the hypometabolic tumor areas.

Figure 4. Identification and spatial mapping of cancer-associated fibroblasts (CAFs) in CSCC. (A) UMAP of 9,836 fibroblasts. CAFs were outlined in red. (B) Violin plot showing the expression levels of *ACTA2, POSTN, ITGB4,* and *FAP* in the 10 clusters of fibroblasts. (C) Box plot showing the stemness and cell cycle scores of fibroblasts, CAFs, and cancer cells. The p-values were determined by the Wilcoxon signed-rank test. (D) GSVA results of fibroblasts, CAFs, and cancer cells. (E) DEGs identified in fibroblasts, CAFs, and cancer cells. (F) Spatially projected CAFs in representative ST slides. The projected area was determined based on the MIA score of CAFs, *POSTN* expression pattern, and the IHC staining results of POSTN (see **Figure S6B&C**). (G) Venn map showing the number of DEGs identified in the three ST samples with both CAFs+ and CAFs- tumors. (H) Bar plot showing the numbers of up- and down-regulated genes in three ST samples. The CAFs+ tumors were compared to the CAFs- tumors. (I) Dot plot of enriched GO terms for up- and down-regulated DEGs identified in H.

Figure 5. Functional analysis of CAFs. (A) Box plot showing the abundance of immune cells in CAFs+ and CAFs- tumors in three samples. The p-values were determined by the Student's t-test. (B) Ligand-receptor communication network between CAFs and different cervical cells predicted by snRNA-seq data. Right, heatmap of the top predicted ligands expressed by CAFs. Middle, heatmap of ligand-

receptor pairs between CAFs and different cell types in CSCC. Bottom, expression heatmap of top receptors regulated by CAFs in different cell types. CECs, columnar epithelial cells; ECM, extracellular matrix. (C) Bar plot showing the integrin types involved in the communication between CAFs and the other cell types. (D) Expression heatmap of gene sets related to functions of CAFs in tumor development in snRNA-seq data. (E) The progression-free and overall survival rates of CSCC patients estimated based on the signature scores of the marker gene set for CAFs. The analysis was conducted using the TCGA dataset. (F) Representative IHC staining patterns of POSTN in the stroma adjacent to tumor areas in FFPE CSCC samples. (G) Box plot showing the association between expression scores (Mean ± SD) of POSTN in CSCC samples and different clinical characteristics (from left to right): tumor differentiation degree, clinical stages, lymph node metastasis, and squamous cell antigen concentration. (H) Schematic summary of the major functions of CAFs in CSCC.

SUPPLEMENTARY INFORMATION

Figure S1. Experimental details and the annotation process of ST slides. (A) Clinical characteristics and experimental details of cervical samples from CSCC and non-cancer patients. (B) UMAP of ST bins from 18 cervical samples. (C) Annotation results of 2 representative ST slides. (D) Expression of tissue-specific genes in 2 representative ST slides. The outline color indicates different tissue types: purple, tumor; blue, stroma with inflammation; green, stroma; red, blood vessel; brown, gland. Violin plots display the gene expression levels in the ST areas.

Figure S2. Annotation results of the 18 ST slides. Non-cancer samples included TJH10, TJH11, and TJH15. All the other samples were CSCC samples. Six genes were selected to represent different tissue types. Cancer cells, *CDKN2A*; Endothelial cells, *PECAM1*; Fibroblasts, *VIM*; Smooth muscle cells, *ADRA2A*; Plasma cells, *IGKC*; Columnar epithelial cells, *MUC5B*.

Figure S3. HPV reads in the ST sequencing data of 15 CSCC tissues. (A) Mapping of HPV reads against the corresponding reference genome. (B) Schematic plot showing the genomic arrangement of the HPV genes in a linear form.

Figure S4. Expression of HPV genes in the tumor areas of selected CSCC ST slides. The annotation result for each sample can be found in **Figure S2**.

Figure S5. Heatmap showing the relationships between tumor differentiation, metabolism, and immunity. Correlation analysis was conducted using 197,181 bins (bin size: 100x100 spots) from 15 CSCC samples. The p-values were determined by the Student's t-test and are indicated by the color and size of the squares.

Figure S6. Characterization of CAFs. (A) Correlation between snRNA-seq cell types and the ST areas defined by multimodal intersection analysis (MIA). (B) Expression of *POSTN* in ST clusters and the associated MIA scores for CAFs. (C) Spatial clustering of CAFs in ST slides and the IHC staining of POSTN in corresponding serial sections of CSCC samples. (D) Spatial prediction of immunocytes in ST slides. (E) The abundance of macrophages (M1, M2, and TAM) in CAFs+ and CAFs- tumors.

Table S1. Clinical characteristics and experimental details for samples.

Table S2. snRNA-seq data statistics.

Table S3. Marker genes for the annotation of cell types in CSCC.

Table S4. ST data statistics.

Table S5. Expression matrix of immune genes in ST areas.

Table S6. DEGs for fibroblasts, CAFs, and cancer cells.

Table S7. DEGs for CAFs+ and CAFs- tumors.

REFERENCES

1. Arbyn, M. *et al*. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *The Lancet Global Health* **8**, e191–e203 (2020).
2. Sung, H. *et al*. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J Clin* **71**, 209–249 (2021).

3. WHO. A Global Strategy for elimination of cervical cancer - PAHO/WHO | Pan American Health Organization. http://www.paho.org/en/topics/cervical-cancer/global-strategy-elimination-cervical-cancer (2018).
4. de Sanjose, S. *et al.* Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *The Lancet Oncology* **11**, 1048–1056 (2010).
5. Doorbar, J. *et al.* The Biology and Life-Cycle of Human Papillomaviruses. *Vaccine* **30**, F55–F70 (2012).
6. Friedlander, M. & Grogan, M. Guidelines for the Treatment of Recurrent and Metastatic Cervical Cancer. *The Oncologist* **7**, 342–347 (2002).
7. Gupta, S. *et al.* Neoadjuvant Chemotherapy Followed by Radical Surgery Versus Concomitant Chemotherapy and Radiotherapy in Patients With Stage IB2, IIA, or IIB Squamous Cervical Cancer: A Randomized Controlled Trial. *JCO* **36**, 1548–1555 (2018).
8. Tewari, K. S. *et al.* Bevacizumab for advanced cervical cancer: final overall survival and adverse event analysis of a randomised, controlled, open-label, phase 3 trial (Gynecologic Oncology Group 240). *The Lancet* **390**, 1654–1663 (2017).
9. Bhatla, N., Aoki, D., Sharma, D. N. & Sankaranarayanan, R. Cancer of the cervix uteri. *Int J Gynecol Obstet* **143**, 22–36 (2018).
10. Tewari, K. S., Java, J. J., Gatcliffe, T. A., Bookman, M. A. & Monk, B. J. Chemotherapy-induced neutropenia as a biomarker of survival in advanced ovarian carcinoma: An exploratory study of the Gynecologic Oncology Group. *Gynecologic Oncology* **133**, 439–445 (2014).
11. Orbegoso, C., Murali, K. & Banerjee, S. The current status of immunotherapy for cervical cancer. *Reports of Practical Oncology & Radiotherapy* **23**, 580–588 (2018).

12. Chung, H. C. *et al.* Efficacy and Safety of Pembrolizumab in Previously Treated Advanced Cervical Cancer: Results From the Phase II KEYNOTE-158 Study. *JCO* **37**, 1470–1478 (2019).
13. Attademo, L. *et al.* Immunotherapy in cervix cancer. *Cancer Treatment Reviews* **90**, 102088 (2020).
14. Frenel, J.-S. *et al.* Safety and Efficacy of Pembrolizumab in Advanced, Programmed Death Ligand 1–Positive Cervical Cancer: Results From the Phase Ib KEYNOTE-028 Trial. *JCO* **35**, 4035–4041 (2017).
15. Santin, A. D. *et al.* Phase II evaluation of nivolumab in the treatment of persistent or recurrent cervical cancer (NCT02257528/NRG-GY002). *Gynecologic Oncology* **157**, 161–166 (2020).
16. Naumann, R. W. *et al.* Safety and Efficacy of Nivolumab Monotherapy in Recurrent or Metastatic Cervical, Vaginal, or Vulvar Carcinoma: Results From the Phase I/II CheckMate 358 Trial. *JCO* **37**, 2825–2834 (2019).
17. Li, C., Guo, L., Li, S. & Hua, K. Single-cell transcriptomics reveals the landscape of intra-tumoral heterogeneity and transcriptional activities of ECs in CC. *Molecular Therapy - Nucleic Acids* **24**, 682–694 (2021).
18. Kato, H. & Torigoe, T. Radioimmunoassay for tumor antigen of human cervical squamous cell carcinoma. *Cancer* **40**, 1621–1628 (1977).
19. Sun, Y., Sheshadri, N. & Zong, W.-X. SERPINB3 and B4: From biochemistry to biology. *Seminars in Cell & Developmental Biology* **62**, 170–177 (2017).
20. Chen, A. *et al.* Large field of view-spatially resolved transcriptomics at nanoscale resolution. 37 (2021).
21. Gulati, G. S. *et al.* Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**, 405–411 (2020).

22. den Boon, J. A. *et al.* Molecular transitions from papillomavirus infection to cervical precancer and cancer: Role of stromal estrogen receptor signaling. *Proc Natl Acad Sci USA* **112**, E3255–E3264 (2015).

23. Muth, S., Klaric, A., Radsak, M., Schild, H. & Probst, H. C. CD27 expression on Treg cells limits immune responses against tumors. *J Mol Med* (2021) doi:10.1007/s00109-021-02116-9.

24. Tanaka, A. & Sakaguchi, S. Regulatory T cells in cancer immunotherapy. *Cell Res* **27**, 109–118 (2017).

25. Vence, L. *et al.* Characterization and Comparison of GITR Expression in Solid Tumors. *Clin Cancer Res* **25**, 6501–6510 (2019).

26. Zhu, M. M. T. *et al.* Evaluation of glucocorticoid-induced TNF receptor (GITR) expression in breast cancer and across multiple tumor types. *Mod Pathol* **33**, 1753–1763 (2020).

27. Wang, M. *et al.* Exosomal LGALS9 in the cerebrospinal fluid of glioblastoma patients suppressed dendritic cell antigen presentation and cytotoxic T-cell immunity. *Cell Death Dis* **11**, 896 (2020).

28. Zhu, C. *et al.* The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity. *Nat Immunol* **6**, 1245–1252 (2005).

29. Riera-Domingo, C. *et al.* Immunity, Hypoxia, and Metabolism–the Ménage à Trois of Cancer: Implications for Immunotherapy. *Physiological Reviews* **100**, 1–102 (2020).

30. Samec, M. *et al.* Flavonoids Targeting HIF-1: Implications on Cancer Metabolism. *Cancers* **13**, 130 (2021).

31. Moncada, R. *et al.* Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol* **38**, 333–342 (2020).

32. Qin, X. *et al.* TGFβ3-mediated induction of Periostin facilitates head and neck cancer growth and is associated with metastasis. *Sci Rep* **6**, 20587 (2016).
33. Yu, B. *et al.* Periostin secreted by cancer-associated fibroblasts promotes cancer stemness in head and neck cancer by activating protein tyrosine kinase 7. *Cell Death Dis* **9**, 1082 (2018).
34. Yue, H. *et al.* Stromal POSTN induced by TGF-β1 facilitates the migration and invasion of ovarian cancer. *Gynecologic Oncology* **160**, 530–538 (2021).
35. Mantovani, A., Marchesi, F., Malesci, A., Laghi, L. & Allavena, P. Tumour-associated macrophages as treatment targets in oncology. *Nat Rev Clin Oncol* **14**, 399–416 (2017).
36. Peacock, J. W. *et al.* SEMA 3C drives cancer growth by transactivating multiple receptor tyrosine kinases via Plexin B1. *EMBO Mol Med* **10**, 219–238 (2018).
37. Zhang, X., Klamer, B., Li, J., Fernandez, S. & Li, L. A pan-cancer study of class-3 semaphorins as therapeutic targets in cancer. *BMC Med Genomics* **13**, 45 (2020).
38. Bao, S. *et al.* Periostin potently promotes metastatic growth of colon cancer by augmenting cell survival via the Akt/PKB pathway. *Cancer Cell* **5**, 329–339 (2004).
39. Zheng, S. *et al.* CXCL6 fuels the growth and metastases of esophageal squamous cell carcinoma cells both in vitro and in vivo through upregulation of PD□L1 via activation of STAT3 pathway. *J Cell Physiol* **236**, 5373–5386 (2021).
40. Blanco-Gómez, A. *et al.* Stromal SNAI2 Is Required for ERBB2 Breast Cancer Progression. *Cancer Res* **80**, 5216–5230 (2020).
41. Cobaleda, C., Pérez-Caro, M., Vicente-Dueñas, C. & Sánchez-García, I. Function of the Zinc-Finger Transcription Factor *SNAI2* in Cancer and Development. *Annu. Rev. Genet.* **41**, 41–61 (2007).
42. Sheng, W. *et al.* LSD1 Ablation Stimulates Anti-tumor Immunity and Enables Checkpoint Blockade. *Cell* **174**, 549-563.e19 (2018).

43. Asem, M. *et al.* Host Wnt5a Potentiates Microenvironmental Regulation of Ovarian Cancer Metastasis. *Cancer Res* **80**, 1156–1170 (2020).

44. Liu, Q. *et al.* Wnt5a-induced M2 polarization of tumor-associated macrophages via IL-10 promotes colorectal cancer progression. *Cell Commun Signal* **18**, 51 (2020).

45. Liu, J. *et al.* An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* **173**, 400-416.e11 (2018).

46. Zheng, Z.-M. Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front Biosci* **11**, 2286 (2006).

47. Choi, Y. J. *et al.* A Phase II, Prospective, Randomized, Multicenter, Open-Label Study of GX-188E, an HPV DNA Vaccine, in Patients with Cervical Intraepithelial Neoplasia 3. *Clin Cancer Res* **26**, 1616–1623 (2020).

48. Fakhr, E., Modic, Ž. & Cid□Arregui, A. Recent developments in immunotherapy of cancers caused by human papillomaviruses. *Immunology* **163**, 33–45 (2021).

49. Trimble, C. L. *et al.* Safety, efficacy, and immunogenicity of VGX-3100, a therapeutic synthetic DNA vaccine targeting human papillomavirus 16 and 18 E6 and E7 proteins for cervical intraepithelial neoplasia 2/3: a randomised, double-blind, placebo-controlled phase 2b trial. *The Lancet* **386**, 2078–2088 (2015).

50. Kamen, D. L. & Tangpricha, V. Vitamin D and molecular actions on the immune system: modulation of innate and autoimmunity. *J Mol Med (Berl)* **88**, 441–450 (2010).

51. Kong, J. *et al.* Novel role of the vitamin D receptor in maintaining the integrity of the intestinal mucosal barrier. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **294**, G208–G216 (2008).

52. Liu, P. T., Krutzik, S. R. & Modlin, R. L. Therapeutic implications of the TLR and VDR partnership. *Trends in Molecular Medicine* **13**, 117–124 (2007).

53. Liu, P. T. *et al.* Toll-Like Receptor Triggering of a Vitamin D-Mediated Human Antimicrobial Response. *Science* **311**, 1770–1773 (2006).

54. Nursyam, E. W., Amin, Z. & Rumende, C. M. The effect of vitamin D as supplementary treatment in patients with moderately advanced pulmonary tuberculous lesion. *Acta Med Indones* **38**, 3–5 (2006).

55. Fahey, L. M., Raff, A. B., Da Silva, D. M. & Kast, W. M. Reversal of Human Papillomavirus-Specific T Cell Immune Suppression through TLR Agonist Treatment of Langerhans Cells Exposed to Human Papillomavirus Type 16. *J Immunol* **182**, 2919–2928 (2009).

56. Garland, S. M. Imiquimod. *Current Opinion in Infectious Diseases* **16**, 85–89 (2003).

57. Giroux, V. *et al.* Long-lived keratin 15+ esophageal progenitor cells contribute to homeostasis and regeneration. *Journal of Clinical Investigation* **127**, 2378–2391 (2017).

58. Han, W., Hu, C., Fan, Z.-J. & Shen, G.-L. Transcript levels of keratin 1/5/6/14/15/16/17 as potential prognostic indicators in melanoma patients. *Sci Rep* **11**, 1023 (2021).

59. Jin, R. Metallothionein 2A expression is associated with cell proliferation in breast cancer. *Carcinogenesis* **23**, 81–86 (2002).

60. Lin, J. *et al.* KRT 15 as a prognostic biomarker is highly expressed in esophageal carcinoma. *Future Oncology* **16**, 1903–1909 (2020).

61. Enwere, E. K. *et al.* Expression of PD-L1 and presence of CD8-positive T cells in pre-treatment specimens of locally advanced cervical cancer. *Mod Pathol* **30**, 577–586 (2017).

62. Mezache, L., Paniccia, B., Nyinawabera, A. & Nuovo, G. J. Enhanced expression of PD L1 in cervical intraepithelial neoplasia and cervical cancers. *Mod Pathol* **28**, 1594–1602 (2015).

63. Reddy, O. L., Shintaku, P. I. & Moatamed, N. A. Programmed death-ligand 1 (PD-L1) is expressed in a significant number of the uterine cervical carcinomas. *Diagn Pathol* **12**, 45 (2017).

64. Lheureux, S. *et al.* Association of Ipilimumab With Safety and Antitumor Activity in Women With Metastatic or Recurrent Human Papillomavirus–Related Cervical Carcinoma. *JAMA Oncol* **4**, e173776 (2018).

65. Daley, D. *et al.* Dectin 1 activation on macrophages by galectin 9 promotes pancreatic carcinoma and peritumoral immune tolerance. *Nat Med* **23**, 556–567 (2017).

66. Nobumoto, A. *et al.* Galectin-9 suppresses tumor metastasis by blocking adhesion to endothelium and extracellular matrices. *Glycobiology* **18**, 735–744 (2008).

67. Munn, D. H. *et al.* GCN2 Kinase in T Cells Mediates Proliferative Arrest and Anergy Induction in Response to Indoleamine 2,3-Dioxygenase. *Immunity* **22**, 633–642 (2005).

68. Low, H.-Y. *et al.* Reciprocal Regulation Between Indoleamine 2,3-Dioxigenase 1 and Notch1 Involved in Radiation Response of Cervical Cancer Stem Cells. *Cancers* **12**, 1547 (2020).

69. Li, X. *et al.* Navigating metabolic pathways to enhance antitumour immunity and immunotherapy. *Nat Rev Clin Oncol* **16**, 425–441 (2019).

70. Biffi, G. & Tuveson, D. A. Diversity and Biology of Cancer-Associated Fibroblasts. *Physiological Reviews* **101**, 147–176 (2021).

71. Chu, T.-Y., Yang, J.-T., Huang, T.-H. & Liu, H.-W. Crosstalk with Cancer-Associated Fibroblasts Increases the Growth and Radiation Survival of Cervical Cancer Cells. *Radiation Research* **181**, 540–547 (2014).

72. Zhang, X. *et al.* Extracellular vesicles-encapsulated microRNA-10a-5p shed from cancer-associated fibroblast facilitates cervical squamous cell carcinoma cell angiogenesis and tumorigenicity via Hedgehog signaling pathway. *Cancer Gene Ther* **28**, 529–542 (2021).

73. Sahai, E. *et al.* A framework for advancing our understanding of cancer-associated fibroblasts. *Nat Rev Cancer* **20**, 174–186 (2020).

74. Guo, W. *et al.* scCancer: a package for automated processing of single-cell RNA-seq data in cancer. *Briefings in Bioinformatics* **22**, bbaa127 (2021).

75. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).

76. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).

77. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381–386 (2014).

78. Zhang, Q. *et al.* Integrative Analysis of Hypoxia-Associated Signature in Pan-Cancer. *iScience* **23**, 101460 (2020).

79. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. *Cell* **160**, 48–61 (2015).

80. Bindea, G. *et al.* Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity* **39**, 782–795 (2013).

81. Cascone, T. *et al.* Increased Tumor Glycolysis Characterizes Immune Resistance to Adoptive T Cell Therapy. *Cell Metabolism* **27**, 977-987.e4 (2018).

82. Baenke, F., Peck, B., Miess, H. & Schulze, A. Hooked on fat: the role of lipid synthesis in cancer metabolism and tumour development. *Dis Model Mech* **6**, 1353–1363 (2013).

83. Mues, C. *et al.* Regulation of Glucose-6-Phosphatase Gene Expression by Insulin and Metformin. *Horm Metab Res* **41**, 730–735 (2009).

84. Alexe, G. *et al.* High Expression of Lymphocyte-Associated Genes in Node-Negative HER2+ Breast Cancers Correlates with Lower Recurrence Rates. *Cancer Res* **67**, 10669–10676 (2007).

85. Heusinkveld, M. & van der Burg, S. H. Identification and manipulation of tumor associated macrophages in human cancers. *J Transl Med* **9**, 216 (2011).

86. Charoentong, P. *et al.* Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Reports* **18**, 248–262 (2017).

87. Stuart, T. *et al*. Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).

88. Malta, T. M. *et al*. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* **173**, 338-354.e15 (2018).

89. Jin, S. *et al*. Inference and analysis of cell-cell communication using CellChat. *Nat Commun* **12**, 1088 (2021).

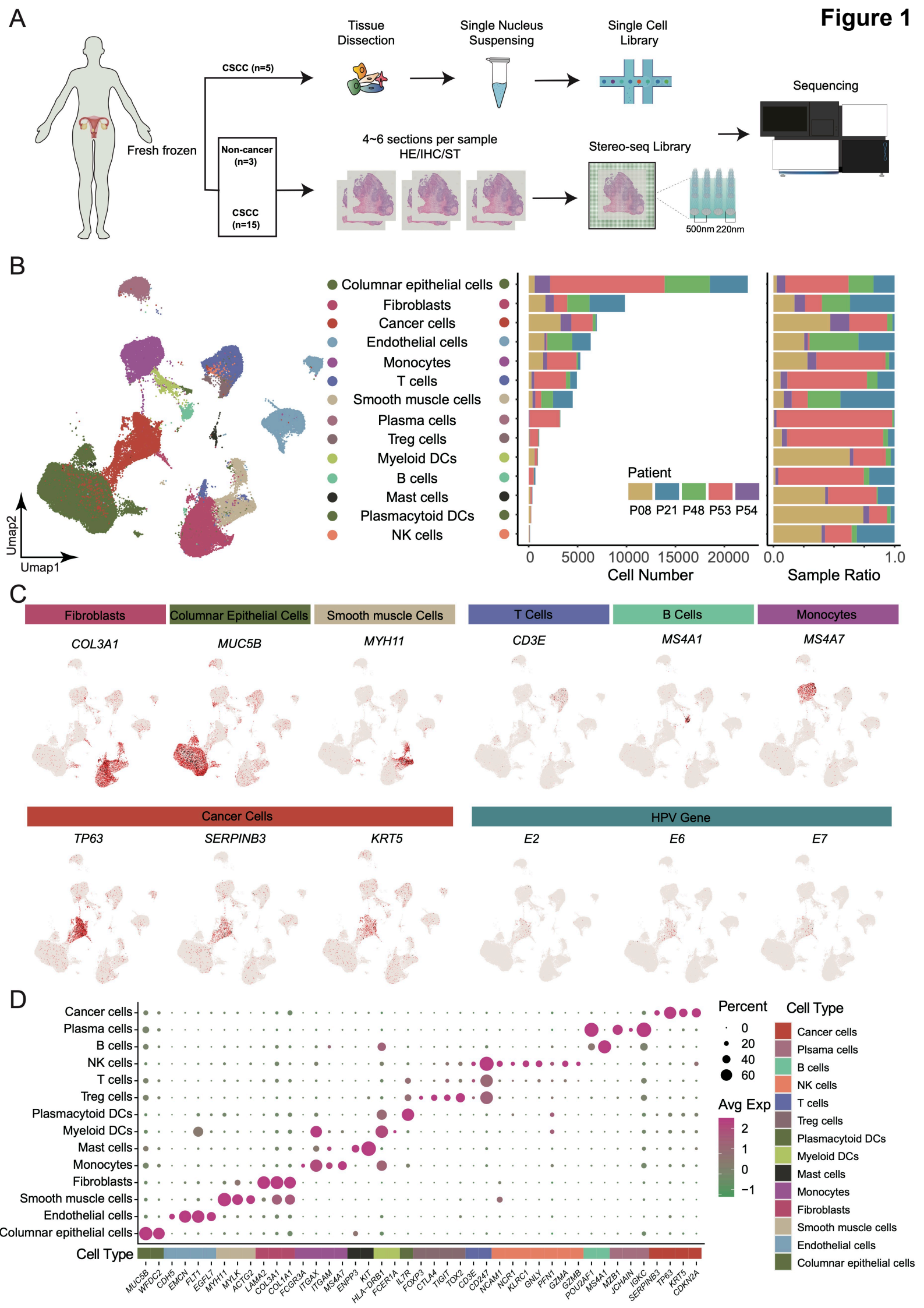


Figure 1
A
Fresh frozen
CSCC (n=5)
Non-cancer (n=3)
CSCC (n=15)
Tissue Dissection
Single Nucleus Suspensing
Single Cell Library
4~6 sections per sample HE/IHC/ST
Stereo-seq Library
500nm 220nm
Sequencing
B
Umap2
Umap1
Columnar epithelial cells
Fibroblasts
Cancer cells
Endothelial cells
Monocytes
T cells
Smooth muscle cells
Plasma cells
Treg cells
Myeloid DCs
B cells
Mast cells
Plasmacytoid DCs
NK cells
Patient
P08 P21 P48 P53 P54
0 5000 10000 15000 20000
Cell Number
0.0 1.0
Sample Ratio
C
Fibroblasts
Columnar Epithelial Cells
Smooth muscle Cells
T Cells
B Cells
Monocytes
COL3A1
MUC5B
MYH11
CD3E
MS4A1
MS4A7
Cancer Cells
HPV Gene
TP63
SERPINB3
KRT5
E2
E6
E7
D
Cancer cells
Plasma cells
B cells
NK cells
T cells
Treg cells
Plasmacytoid DCs
Myeloid DCs
Mast cells
Monocytes
Fibroblasts
Smooth muscle cells
Endothelial cells
Columnar epithelial cells
Cell Type
MUC5B WFDC2 CDH6 EMCN FLT1 EGFL7 MYH11 MYLK ACTG2 LAMA2 COL3A1 COL1A1 FCGR3A ITGAX ITGAM MS4A7 ENPP3 KIT HLA-DRB1 FCER1A IL7R FOXP3 CTLA4 TIGIT TOX2 CD3E CD247 NCAM1 NCR1 KLRC1 GNLY PFN1 GZMA GZMB POU2AF1 MS4A1 MZB1 JCHAIN IGKC SERPINB3 TP63 KRT5 CDKN2A
Percent
0
20
40
60
Avg Exp
2
1
0
−1
Cell Type
Cancer cells
Plsama cells
B cells
NK cells
T cells
Treg cells
Plasmacytoid DCs
Myeloid DCs
Mast cells
Monocytes
Fibroblasts
Smooth muscle cells
Endothelial cells
Columnar epithelial cells

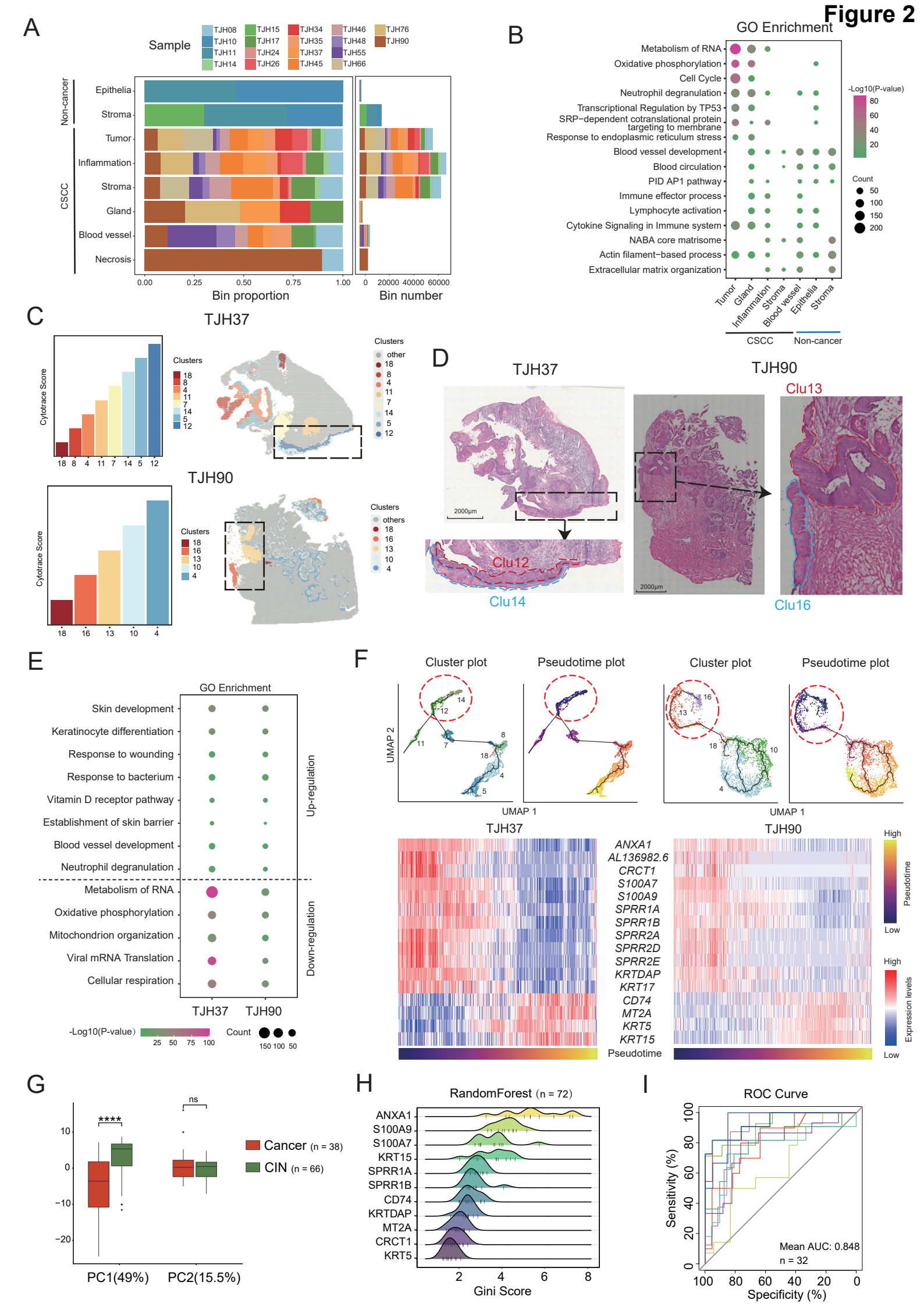


Figure 2
A
Sample
TJH08
TJH10
TJH11
TJH14
TJH15
TJH17
TJH24
TJH26
TJH34
TJH35
TJH37
TJH45
TJH46
TJH48
TJH55
TJH66
TJH76
TJH90
Non-cancer
Epithelia
Stroma
CSCC
Tumor
Inflammation
Stroma
Gland
Blood vessel
Necrosis
Bin proportion
Bin number
B
GO Enrichment
Metabolism of RNA
Oxidative phosphorylation
Cell Cycle
Neutrophil degranulation
Transcriptional Regulation by TP53
SRP−dependent cotranslational protein targeting to membrane
Response to endoplasmic reticulum stress
Blood vessel development
Blood circulation
PID AP1 pathway
Immune effector process
Lymphocyte activation
Cytokine Signaling in Immune system
NABA core matrisome
Actin filament−based process
Extracellular matrix organization
-Log10(P-value)
Count
CSCC
Non-cancer
C
TJH37
Cytotrace Score
Clusters
TJH90
D
TJH37
TJH90
Clu13
Clu12
Clu14
Clu16
2000μm
E
GO Enrichment
Skin development
Keratinocyte differentiation
Response to wounding
Response to bacterium
Vitamin D receptor pathway
Establishment of skin barrier
Blood vessel development
Neutrophil degranulation
Metabolism of RNA
Oxidative phosphorylation
Mitochondrion organization
Viral mRNA Translation
Cellular respiration
Up-regulation
Down-regulation
TJH37
TJH90
-Log10(P-value)
Count
F
Cluster plot
Pseudotime plot
UMAP 2
UMAP 1
TJH37
TJH90
ANXA1
AL136982.6
CRCT1
S100A7
S100A9
SPRR1A
SPRR1B
SPRR2A
SPRR2D
SPRR2E
KRTDAP
KRT17
CD74
MT2A
KRT5
KRT15
Pseudotime
High
Low
Expression levels
G
****
ns
Cancer (n = 38)
CIN (n = 66)
PC1(49%)
PC2(15.5%)
H
RandomForest (n = 72)
ANXA1
S100A9
S100A7
KRT15
SPRR1A
SPRR1B
CD74
KRTDAP
MT2A
CRCT1
KRT5
Gini Score
I
ROC Curve
Sensitivity (%)
Specificity (%)
Mean AUC: 0.848
n = 32

Figure 3

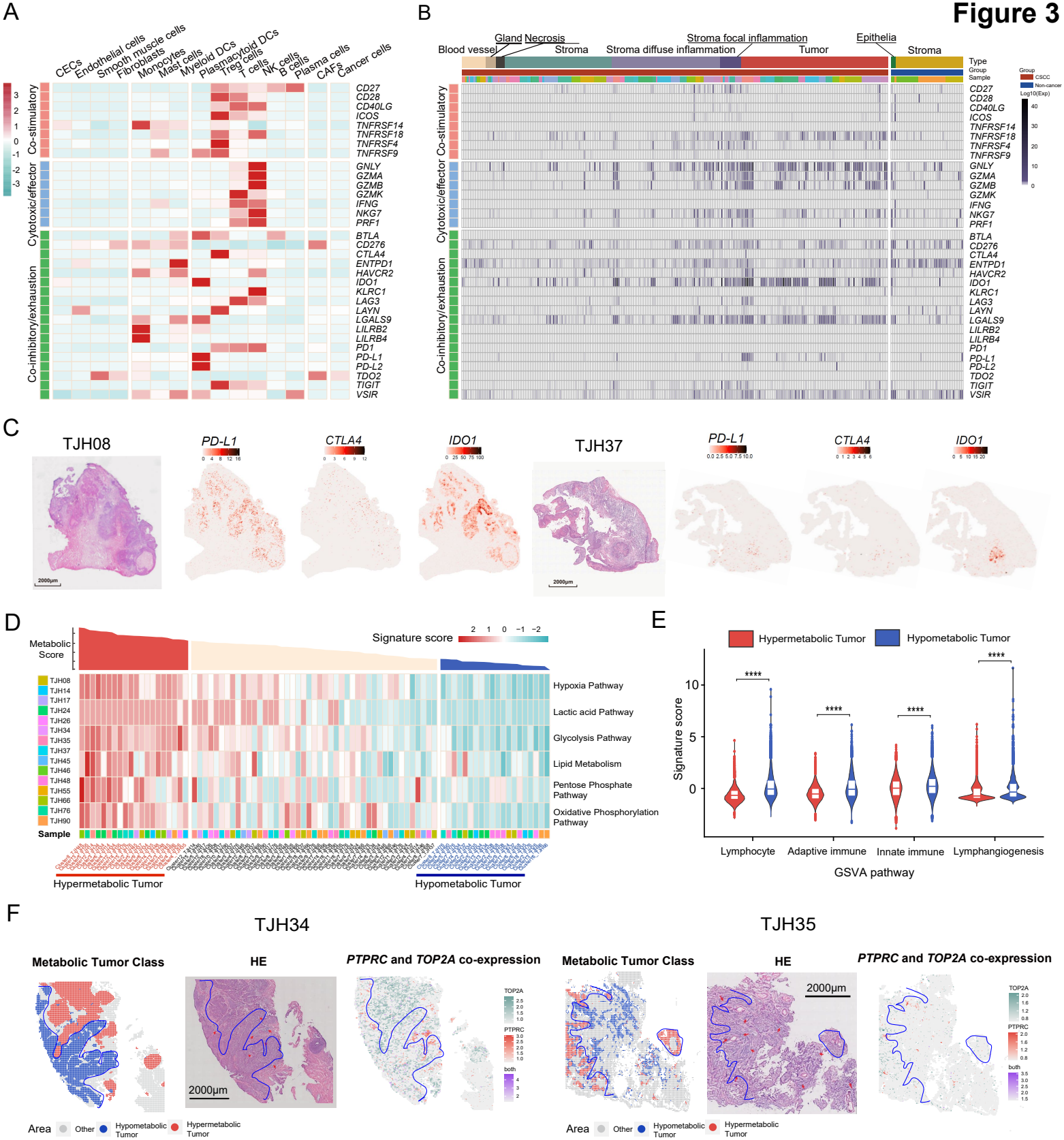

A
CECs
Endothelial cells
Smooth muscle cells
Fibroblasts
Monocytes
Mast cells
Myeloid DCs
Plasmacytoid DCs
Treg cells
T cells
NK cells
B cells
Plasma cells
CAFs
Cancer cells
Co-stimulatory
Cytotoxic/effector
Co-inhibitory/exhaustion
CD27
CD28
CD40LG
ICOS
TNFRSF14
TNFRSF18
TNFRSF4
TNFRSF9
GNLY
GZMA
GZMB
GZMK
IFNG
NKG7
PRF1
BTLA
CD276
CTLA4
ENTPD1
HAVCR2
IDO1
KLRC1
LAG3
LAYN
LGALS9
LILRB2
LILRB4
PD1
PD-L1
PD-L2
TDO2
TIGIT
VSIR
B
Blood vessel
Gland Necrosis
Stroma
Stroma diffuse inflammation
Stroma focal inflammation
Tumor
Epithelia
Stroma
Type
Group
Sample
Log10(Exp)
C
TJH08
PD-L1
CTLA4
IDO1
TJH37
PD-L1
CTLA4
IDO1
2000μm
D
Metabolic Score
Signature score
Hypoxia Pathway
Lactic acid Pathway
Glycolysis Pathway
Lipid Metabolism
Pentose Phosphate Pathway
Oxidative Phosphorylation Pathway
Sample
Hypermetabolic Tumor
Hypometabolic Tumor
E
Hypermetabolic Tumor
Hypometabolic Tumor
Signature score
Lymphocyte
Adaptive immune
Innate immune
Lymphangiogenesis
GSVA pathway
F
TJH34
TJH35
Metabolic Tumor Class
HE
PTPRC and TOP2A co-expression
2000μm
Area
Other
Hypometabolic Tumor
Hypermetabolic Tumor

Figure 4

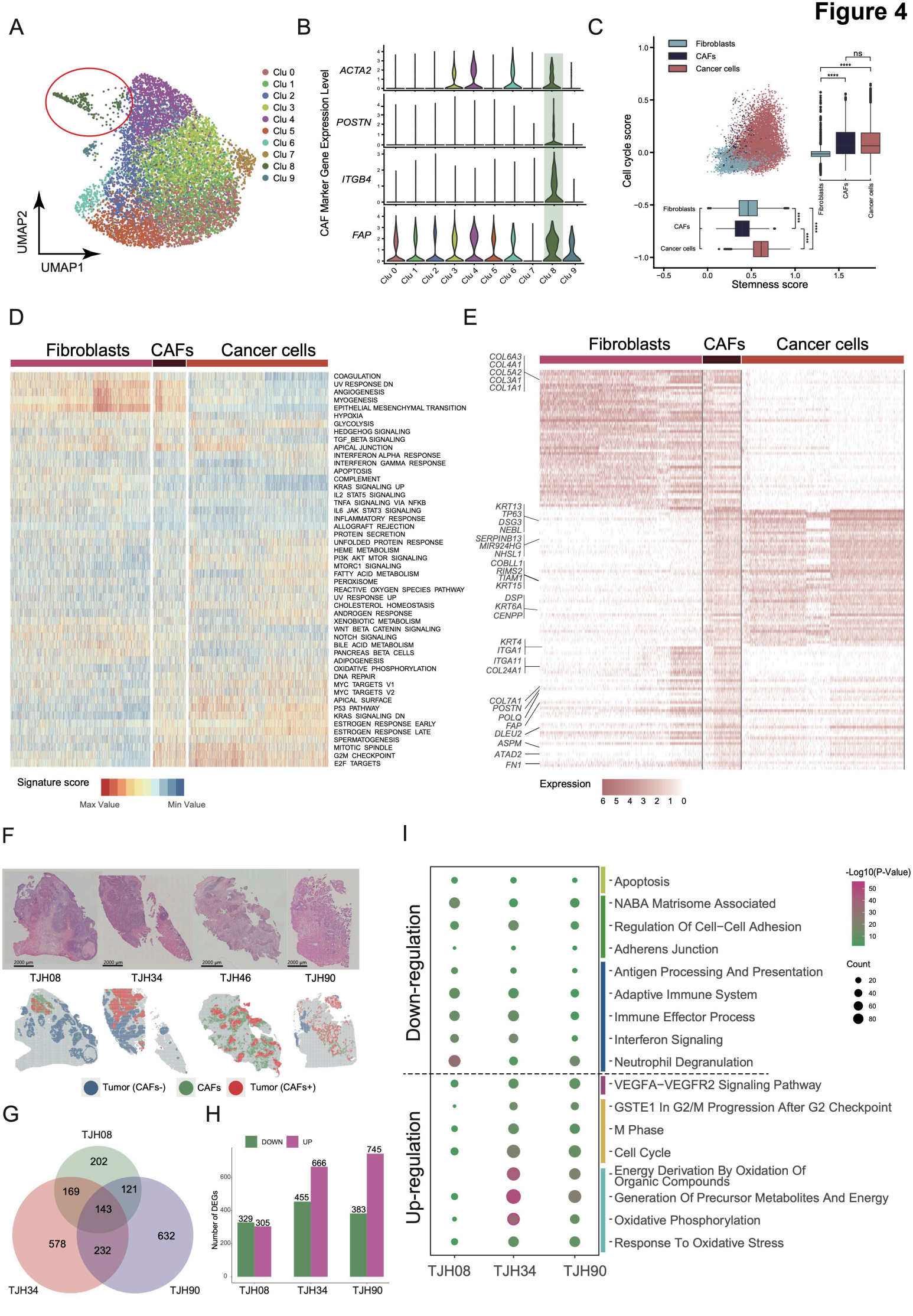
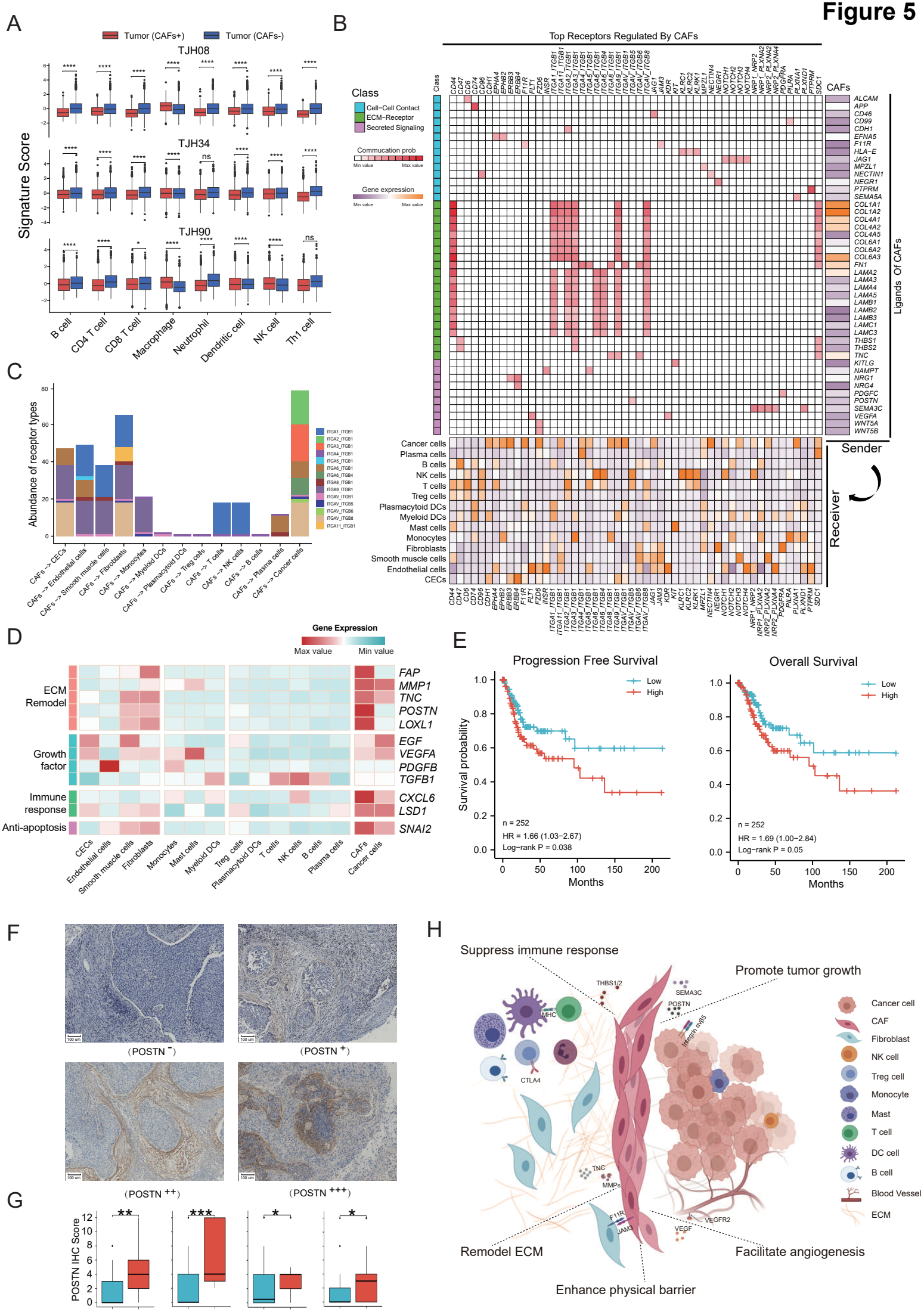

A
Clu 0
Clu 1
Clu 2
Clu 3
Clu 4
Clu 5
Clu 6
Clu 7
Clu 8
Clu 9
UMAP2
UMAP1
B
CAF Marker Gene Expression Level
ACTA2
POSTN
ITGB4
FAP
Clu 0
Clu 1
Clu 2
Clu 3
Clu 4
Clu 5
Clu 6
Clu 7
Clu 8
Clu 9
C
Fibroblasts
CAFs
Cancer cells
Cell cycle score
Stemness score
ns
****
D
Fibroblasts
CAFs
Cancer cells
COAGULATION
UV RESPONSE DN
ANGIOGENESIS
MYOGENESIS
EPITHELIAL MESENCHYMAL TRANSITION
HYPOXIA
GLYCOLYSIS
HEDGEHOG SIGNALING
TGF_BETA SIGNALING
APICAL JUNCTION
INTERFERON ALPHA RESPONSE
INTERFERON GAMMA RESPONSE
APOPTOSIS
COMPLEMENT
KRAS SIGNALING UP
IL2 STAT5 SIGNALING
TNFA SIGNALING VIA NFKB
IL6 JAK STAT3 SIGNALING
INFLAMMATORY RESPONSE
ALLOGRAFT REJECTION
PROTEIN SECRETION
UNFOLDED PROTEIN RESPONSE
HEME METABOLISM
PI3K AKT MTOR SIGNALING
MTORC1 SIGNALING
FATTY ACID METABOLISM
PEROXISOME
REACTIVE OXYGEN SPECIES PATHWAY
UV RESPONSE UP
CHOLESTEROL HOMEOSTASIS
ANDROGEN RESPONSE
XENOBIOTIC METABOLISM
WNT BETA CATENIN SIGNALING
NOTCH SIGNALING
BILE ACID METABOLISM
PANCREAS BETA CELLS
ADIPOGENESIS
OXIDATIVE PHOSPHORYLATION
DNA REPAIR
MYC TARGETS V1
MYC TARGETS V2
APICAL SURFACE
P53 PATHWAY
KRAS SIGNALING DN
ESTROGEN RESPONSE EARLY
ESTROGEN RESPONSE LATE
SPERMATOGENESIS
MITOTIC SPINDLE
G2M CHECKPOINT
E2F TARGETS
Signature score
Max Value
Min Value
E
Fibroblasts
CAFs
Cancer cells
COL6A3
COL4A1
COL5A2
COL3A1
COL1A1
KRT13
TP63
DSG3
NEBL
SERPINB13
MIR924HG
NHSL1
COBLL1
RIMS2
TIAM1
KRT15
DSP
KRT6A
CENPP
KRT4
ITGA1
ITGA11
COL24A1
COL7A1
POSTN
POLQ
FAP
DLEU2
ASPM
ATAD2
FN1
Expression
6 5 4 3 2 1 0
F
TJH08
TJH34
TJH46
TJH90
Tumor (CAFs-)
CAFs
Tumor (CAFs+)
G
TJH08
202
169
121
143
578
232
632
TJH34
TJH90
H
DOWN
UP
Number of DEGs
329
305
455
666
383
745
TJH08
TJH34
TJH90
I
Down-regulation
Up-regulation
Apoptosis
NABA Matrisome Associated
Regulation Of Cell−Cell Adhesion
Adherens Junction
Antigen Processing And Presentation
Adaptive Immune System
Immune Effector Process
Interferon Signaling
Neutrophil Degranulation
VEGFA−VEGFR2 Signaling Pathway
GSTE1 In G2/M Progression After G2 Checkpoint
M Phase
Cell Cycle
Energy Derivation By Oxidation Of Organic Compounds
Generation Of Precursor Metabolites And Energy
Oxidative Phosphorylation
Response To Oxidative Stress
TJH08
TJH34
TJH90
-Log10(P-Value)
Count

Figure 5


A
Tumor (CAFs+)
Tumor (CAFs-)
TJH08
TJH34
TJH90
Signature Score
B cell
CD4 T cell
CD8 T cell
Macrophage
Neutrophil
Dendritic cell
NK cell
Th1 cell
B
Top Receptors Regulated By CAFs
Class
Cell−Cell Contact
ECM−Receptor
Secreted Signaling
Commucation prob
Gene expression
CAFs
Ligands Of CAFs
Sender
Receiver
Cancer cells
Plasma cells
B cells
NK cells
T cells
Treg cells
Plasmacytoid DCs
Myeloid DCs
Mast cells
Monocytes
Fibroblasts
Smooth muscle cells
Endothelial cells
CECs
C
Abundance of receptor types
CAFs -> CECs
CAFs -> Endothelial cells
CAFs -> Smooth muscle cells
CAFs -> Fibroblasts
CAFs -> Monocytes
CAFs -> Myeloid DCs
CAFs -> Plasmacytoid DCs
CAFs -> Treg cells
CAFs -> T cells
CAFs -> NK cells
CAFs -> B cells
CAFs -> Plasma cells
CAFs -> Cancer cells
D
Gene Expression
Max value
Min value
ECM Remodel
Growth factor
Immune response
Anti-apoptosis
FAP
MMP1
TNC
POSTN
LOXL1
EGF
VEGFA
PDGFB
TGFB1
CXCL6
LSD1
SNAI2
E
Progression Free Survival
Overall Survival
Low
High
Survival probability
Months
n = 252
HR = 1.66 (1.03−2.67)
Log−rank P = 0.038
HR = 1.69 (1.00−2.84)
Log−rank P = 0.05
F
(POSTN ⁻)
(POSTN ⁺)
(POSTN ⁺⁺)
(POSTN ⁺⁺⁺)
G
POSTN IHC Score
Well/Moderate
Low
Stage I-II
Stage III
N0
N1+
SCC <=1.5
SCC >1.5
H
Suppress immune response
Promote tumor growth
Remodel ECM
Facilitate angiogenesis
Enhance physical barrier
THBS1/2
SEMA3C
POSTN
MHC
CTLA4
TNC
MMPs
F11R
JAM3
VEGF
VEGFR2
Cancer cell
CAF
Fibroblast
NK cell
Treg cell
Monocyte
Mast
T cell
DC cell
B cell
Blood Vessel
ECM